

## **Relación entre los patrones de búsqueda de internet y los movimientos del Mercado Integrado Latinoamericano**

**Mariana Giraldo Moreno**  
**Luisa Toro Moreno**

**Ingeniera Administradora**  
**Ingeniera Financiera**

**Juan Pablo Valencia Arango**

**Msc Estadística**



**UNIVERSIDAD EIA**  
**INGENIERÍA ADMINISTRATIVA-INGENIERÍA FINANCIERA**  
**ENVIGADO**  
**2016**

## CONTENIDO

### [INTRODUCCIÓN](#)

#### [1. PRELIMINARES](#)

##### [1.1 Planteamiento del problema](#)

###### [1.1.1 Contexto y caracterización del problema](#)

###### [1.1.2 Formulación del problema](#)

##### [1.2 Objetivos del proyecto](#)

###### [1.2.1 Objetivo General](#)

###### [1.2.2 Objetivos Específicos](#)

##### [1.3 Marco de referencia](#)

###### [1.3.1 Antecedentes](#)

###### [1.3.2 Marco Teórico](#)

##### [1.4 Hipótesis](#)

#### [2. METODOLOGÍA](#)

#### [3. DESARROLLO DEL PROYECTO](#)

##### [3.1 Objetivo 1](#)

###### [3.1.1 Determinar el índice representativo del mercado de cada país del MILA](#)

###### [3.1.2 Definir los activos que se utilizarán](#)

###### [3.1.3 Establecer los artículos de Wikipedia para cada activo seleccionado](#)

###### [3.1.4 Definir las palabras clave asociadas a cada activo](#)

##### [3.2 Objetivo 2](#)

###### [3.2.1 Obtener los datos de los activos](#)

**Relación entre los patrones de búsqueda de internet y los movimientos del MILA**

[3.2.2 Desarrollar el código en R para la implementación de la metodología](#)

[3.3 Objetivo 3](#)

[3.3.1 El resultado de las pruebas realizadas con su respectivo análisis.](#)

[3.3.2 Conclusiones generales sobre las relaciones entre los patrones de búsqueda de internet y los movimientos en el mercado seleccionado.](#)

[4. DISCUSIÓN DE LOS RESULTADOS](#)

[5. CONCLUSIONES Y CONSIDERACIONES FINALES](#)

[6. MATERIAL COMPLEMENTARIO](#)

[BIBLIOGRAFÍA](#)

[ANEXO 1](#)

[ANEXO 2](#)

## LISTA DE TABLAS

Tabla 1: Uso del Internet

Tabla 2: Participación en total motores de búsqueda a nivel global de computadores

Tabla 3: Índices representativos por país

Tabla 4: Activos Colombia

Tabla 5: Activos México

Tabla 6: Activos de Chile

Tabla 7: Activos Perú

Tabla 8: Retornos Acumulados por Estrategia y País

Tabla 9: Resultados de Pruebas Estadísticas

## LISTA DE FIGURAS

Imagen 1:  $\Delta t$  Wikipedia

Imagen 2:  $\Delta t$  Google

Imagen 3: Retornos acumulados Éxito

Imagen 4: Retornos acumulados Colombia

Imagen 5: Retornos acumulados México

Imagen 6: Retornos acumulados Chile

Imagen 7: Retornos acumulados Perú

## GLOSARIO

**POSICIÓN CORTA:** cuando se piensa que el valor de un activo caerá a corto o mediano plazo el inversionista vende el activo con la intención de recomprarlo cuando el precio del activo baje. Esto con el objetivo de obtener una ganancia cuando el precio baje (Harvey, 2011).

**POSICIÓN LARGA:** cuando se piensa que el valor de un activo subirá a corto o mediano plazo el inversionista compra el activo con la intención de venderlo cuando el precio del activo suba. Esto con el objetivo de obtener una ganancia (Harvey, Nasdaq, 2011).

**RENTA VARIABLE:** el mercado de renta variables es aquel en el que los activos que se transan no tienen preestablecido el rendimiento. La rentabilidad está ligada a las utilidades obtenidas por la empresa en la cual se invierte, así como a las posibles variaciones en los precios de los valores, dadas por las condiciones existentes en el mercado. El título característico de este mercado son las acciones (BVC, 2015a).

**RETORNO:** es la ganancia o pérdida de un título en un periodo de tiempo. El retorno consiste en el ingreso y ganancia de capital relativa en una inversión. Normalmente es presentado como un porcentaje (Investopedia, 2015a).

**VOLATILIDAD:** es la fluctuación o la amplitud de los movimientos respecto a la media de un activo en un periodo de tiempo. La volatilidad trata de cuantificar la variabilidad o dispersión de un activo respecto a su tendencia central (BBVA, 2015).

Ayuda a medir la incertidumbre de un mercado y se utiliza principalmente como medida del riesgo del activo: cuanto más volátil sea un activo, mayor riesgo tendrá (BBVA, 2015).

**VOLUMEN DE NEGOCIACIÓN:** es el número de acciones o contratos negociados en un mercado durante un periodo de tiempo determinado. Es la cantidad de acciones que cambia de mano de vendedores a la de compradores (Investopedia, 2015b).

Como hay un vendedor por cada comprador, el volumen puede pensarse como la mitad del número de acciones negociadas (NASDAQ, 2015).

## RESUMEN

Las personas comienzan su proceso de toma de decisiones con un intento de recolectar información (Simon, 1957) y en la actualidad la mayor parte de este proceso se lleva a cabo a través de la internet. El análisis de la información de búsquedas realizada en internet podría proporcionar una nueva perspectiva sobre los cambios en el proceso de recolección de información de los inversionistas y los movimientos futuros de los mercados financieros.

Este proyecto pretende utilizar información del tráfico de la página Wikipedia y el buscador de Google para cuantificar la posible relación existente entre los patrones de búsqueda en Google y Wikipedia de algunas empresas listadas en bolsa y el comportamiento de estas en el Mercado Integrado Latinoamericano (MILA) de acuerdo a la metodología Moat (2013)

Para desarrollar el proyecto se seleccionaron los activos listados en bolsa, los artículos de Wikipedia y las palabras clave de búsqueda en Google que se emplearán para el estudio. Luego se implementó una estrategia hipotética de inversión que utilizó esta información como criterios de decisión para comprar o vender los activos y por último se validaron estadísticamente los resultados de la estrategia implementada.

La mayoría de los retornos obtenidos mediante la estrategia implementada no fueron estadísticamente mayores que cero, por lo tanto se concluyó que no se puede afirmar que exista la relación propuesta por Moat(2013) entre las búsquedas en *Google* y *Wikipedia* y los movimientos del mercado accionario de los activos pertenecientes al MILA .

Palabras clave: Mercados Financieros, [modelamiento estadístico](#), Google, Wikipedia, MILA.

## ABSTRACT

People start their decision-making process trying to collect information and today most of this process takes place through the Internet (Simon, 1957). The data analysis of online searches could provide a new perspective on the changes in the process of gathering information from investors and on future movements of financial markets.

This project aims to use the traffic information on the Wikipedia page and on Google search engine to determine whether there is any relationship between search patterns on Google and Wikipedia of some companies listed on the stock market and the behavior of these stocks in the Latin America Integrated Market.

To develop this project, the stocks listed in the stock market, the Wikipedia articles and the search keywords in Google will be selected. Subsequently, a hypothetical investment strategy that uses this information as decision criteria will be implemented, and finally the results from the implemented strategy will be statistically validated.

Most of the returns obtained by the strategy implemented were not statistically greater than zero, therefore we concluded that it can not be said that there is the relationship proposed by Moat (2013) between Google and Wikipedia searches and the movements of stock market assets of the MILA market.

Key words: Financial Markets, Big Data, Google, Wikipedia, MILA



## INTRODUCCIÓN

Predecir el comportamiento de la bolsa es [uno de los retos que enfrenta un inversionista](#). Existen toda clase de estudios basados en análisis técnicos y fundamentales que buscan proveer herramientas que le permitan al inversionista anticiparse a movimientos del mercado. Sin embargo, esto continúa siendo un gran desafío.

Anteriormente, se han realizado estudios para encontrar relaciones entre los patrones de búsqueda de internet y el comportamiento del mercado financiero ([Preis, Reith, & Stanley, 2010](#)) ([Bank, Larch, & Peter, 2011](#)) ([Bordino et al., 2012](#)) ([Zheludev, 2014](#)) ([Moat et al., 2013](#)). No obstante, para el mercado latinoamericano no se han encontrado análisis en este aspecto hasta el momento.

En el presente trabajo se pretende utilizar la información sobre búsquedas en internet, específicamente en Google y Wikipedia, de activos financieros de los países pertenecientes al MILA, es decir, México, Perú, Chile y Colombia, para determinar si existe alguna relación entre los patrones de búsqueda en internet y los movimientos en el mercado financiero.

Inicialmente se presentarán algunos antecedentes de estudios similares, realizados en otros países. Luego, en el marco teórico, se definirán conceptos importantes [que permitan tener total entendimiento de la problemática que se aborda](#). En la metodología se explicarán las técnicas utilizadas para implementar el estudio y se dará un detalle de las actividades necesarias para el cumplimiento de los objetivos.

Más adelante se explicará detalladamente el proceso de obtención de los datos y elección de los activos, la determinación e implementación de la estrategia utilizada, el análisis estadístico de los resultados y por último las conclusiones [y recomendaciones](#) del estudio.

## 1. PRELIMINARES

### 1.1 PLANTEAMIENTO DEL PROBLEMA

#### 1.1.1 Contexto y caracterización del problema

Actualmente vivimos en una sociedad computarizada donde muchas de nuestras acciones dejan rastros digitales. Estos han permitido que diferentes instancias tengan acceso a grandes bases de datos que reflejan las actividades de millones de personas en internet y a utilizarlas para analizar su comportamiento a través del descubrimiento de patrones y conexiones.

Los mercados financieros proveen una gran cantidad de registros detallados sobre las transacciones que se realizan y estos se han utilizado para predecir los movimientos del mercado y tomar mejores decisiones. Sin embargo, los registros reflejan únicamente la decisión final tomada por los inversionistas, es decir, la de comprar o vender un determinado activo, que se da una vez terminado el proceso de decisión.

Los inversionistas comienzan su procesos de toma de decisiones con un intento de recolectar información (Simon, 1957) y en la actualidad la mayor parte de este proceso se lleva a cabo vía internet. El análisis de la información de búsquedas realizada en internet podría proporcionar una nueva perspectiva sobre los cambios en el proceso de recolección de información y los movimientos futuros de los mercados financieros.

Se han realizado varios estudios ((Preis, Reith, & Stanley, 2010), (Bordino et al., 2012), (Bank, Larch, & Peter, 2011), (Alanyali, Moat, & Preis, 2013), (Moat et al., 2013)) que sugieren la relación entre las búsquedas en internet y el comportamiento de los mercados financieros indicando que estas pueden proveer una buena medida del proceso de recolección de información de los inversionistas. Sin embargo, en Latinoamérica y específicamente Colombia no se han encontrado estudios al respecto.

#### 1.1.2 Formulación del problema

No se han encontrado estudios que muestren evidencia de una relación entre las búsquedas en internet y el comportamiento del mercado accionario en Latinoamérica. Este proyecto pretende utilizar información del tráfico de la página Wikipedia y el buscador de Google para encontrar esa relación respondiendo a la pregunta: ¿Existe la relación propuesta por Moat (2013) y planteada en la

hipótesis entre los patrones de búsqueda de palabras claves en *Google* y artículos en *Wikipedia* de algunas empresas listadas en bolsa y el comportamiento del mercado accionario en los países perteneciente al Mercado Integrado Latinoamericano?

## **1.2 OBJETIVOS DEL PROYECTO**

### **1.2.1 Objetivo General**

Cuantificar la relación existente entre las búsquedas de los artículos en Wikipedia y palabras claves en Google de algunos activos listados en bolsa y el comportamiento de estos en el Mercado Integrado Latinoamericano.

### **1.2.2 Objetivos Específicos**

1. Seleccionar los activos listados en bolsa, los artículos de Wikipedia y las palabras claves de búsqueda en Google que se emplearán para el estudio.
2. Implementar una estrategia hipotética de inversión con la información obtenida previamente.
3. Validar estadísticamente los resultados de la estrategia implementada anteriormente.

## **1.3 MARCO DE REFERENCIA**

### **1.3.1 Antecedentes**

Estudios previos han buscado relaciones entre los datos obtenidos a través de internet y el comportamiento de los mercados financieros. Estos se han realizado a partir de información obtenida de diversas fuentes, principalmente a través de Google Trends (Preis, Reith, & Stanley, 2010) (Bank, Larch, & Peter, 2011) aunque también se han utilizado registros de Yahoo (Bordino et al., 2012), Twitter (Zheludev, 2014) y Wikipedia (Moat et al., 2013). Pero la mayoría tiene en común que han sido aplicados al mercado de Estados Unidos.

En el 2010 se realizó un estudio donde se utilizaron los datos históricos del volumen semanal de búsquedas en *Google* para los nombres de las compañías del índice S&P 500 y se encontró que estos están correlacionados con el volumen transado de sus acciones. Además se implementaron métodos para cuantificar correlaciones complejas en series de tiempo que mostraron patrones recurrentes

entre las series de tiempo de los volúmenes de búsqueda y la de los volúmenes transados (Preis, Reith, & Stanley, 2010).

En el año 2011 Bank, Larch, & Peter publicaron un artículo donde se tomaron los datos del volumen de búsquedas de Google para los nombres de las compañías que se transan en la plataforma electrónica de negociación de la Bolsa Alemana Xetra y se encontró que un incremento en el volumen de búsquedas está asociado con un incremento en el volumen de negociación de la acción y con una mayor liquidez. Además, encontraron que un aumento en el volumen de búsquedas está asociado a un incremento en los retornos futuros de corto plazo (Bank, Larch, & Peter, 2011).

Luego en el año 2012 se llevó a cabo un estudio similar, pero en este caso se trabajó con los datos del volumen diario de búsquedas en Yahoo de los nombres de las empresas inscritas en el índice NASDAQ-100. Este estudio demostró que estas búsquedas están correlacionadas con los volúmenes diarios de negociación de las acciones respectivas (Bordino et al., 2012).

Más adelante, en el 2013, los investigadores Preis, Moat y Stanley analizaron las ediciones de Financial Times desde enero 2 de 2007 hasta el diciembre 31 de 2012 y encontraron una correlación positiva entre número diario de menciones en el Financial Times y el volumen diario de transacciones en la bolsa para las empresas estudiadas (Alanyali, Moat, & Preis, 2013).

También se han realizado estudios utilizando los datos obtenidos de redes sociales para comprender los movimientos del mercado. En el 2014 Ilya Zheludev, Robert Smith & Tomaso Aste sugirieron que el uso de técnicas de análisis de sentimiento sobre tweets que hacen referencia a acciones pertenecientes al índice S&P500 puede proveer información sobre los precios futuros de dichas acciones en un periodo de horas, sin embargo, esta relación solo se encontró para determinadas acciones.

En el 2013 Tobias Preis, Susannah Moat y Eugene Stanley realizaron un estudio donde implementaron una estrategia hipotética de inversión sobre el Dow Jones Industrial Average DJIA basada en el volumen de búsquedas de 98 términos relacionados con el mercado accionario los cuales fueron obtenidos con la herramienta Google Trends. Ellos encontraron que incrementos en el número de búsquedas de estas palabras claves están relacionados con futuras caídas en el mercado accionario (Preis, Moat, & Stanley, 2013).

En el 2013 los investigadores Susannah Moat, Chester Curme, Dror Y. Kenett, Eugene Stanley y Tobias Preis realizaron un estudio donde utilizaron los datos

históricos artículos financieros de la enciclopedia en línea Wikipedia y en base a estos desarrollaron una estrategia hipotética de inversión sobre el Índice Dow Jones Industrial Average. Estos encontraron evidencia de que el cambio en el número de visitas a páginas de temas financieros en Wikipedia pueden contener información que permita conocer comportamientos posteriores en el mercado accionario (Moat et al., 2013).

Otros estudios han analizado la relación de las búsquedas en internet con factores como la actividad económica en una industria.

Hal Varian y Hyunyoung Choi, economistas de Google, en su estudio “Predicting the present with Google Trends” utilizaron las búsquedas de algunos términos financieros y llegaron a la conclusión de que los datos de búsqueda están correlacionados con la actual actividad económica en una industria. También sugirieron que esto puede servir para predecir las publicaciones futuras de datos en estas industrias (Choi & Varian, 2012).

En el 2009 también se analizó la relación de las búsquedas en Google con el mercado de viviendas y bienes raíz en Estados Unidos y, no solo se logró predecir la tendencia en este mercado sino que se halló que un índice de búsqueda de viviendas pronostica el futuro mercado de estas (Wu & Brynjolfsson, 2009).

En Colombia han sido utilizados los patrones de búsqueda de Google para predecir indicadores macroeconómicos como el PIB. El Ministerio de Hacienda, basándose en la metodología de pronóstico de corto plazo de series económicas diseñada por Choi y Varian (2011), construyó unos Indicadores Sectoriales Adelantados de Actividad (ISAAC) con los cuales permiten anticipar las tendencias del PIB de las nueve grandes ramas de actividad nacional en el corto plazo en Colombia (MinHacienda, 2013).

Este estudio también se basó en las investigaciones realizadas en Estados Unidos por Francesco D’Amuri y Juri Marcucci en el 2009, que desencadenaron en la creación de un indicador que intenta predecir el desempleo en USA utilizando el índice de búsquedas de empleo de google y el IC (Francesco & Marcucci, 2009).

### **1.3.2 Marco Teórico**

#### **1.3.2.1 Mercado de valores**

##### **Mercado Integrado Latinoamericano**

#### Relación entre los patrones de búsqueda de internet y los movimientos del MILA

El Mercado Integrado Latinoamericano MILA es un mercado regional para la negociación de títulos de renta variable que surgió como resultado del acuerdo firmado entre la Bolsa de Comercio de Santiago, la Bolsa de Valores de Colombia y la Bolsa de Valores de Lima, así como de los depósitos Deceval, DCV y Cavali, las cuales, desde 2009, iniciaron este proceso (MILA, 2015).

En el 2011 el MILA entró en operación para inversionistas e intermediarios de Chile, Colombia y Perú y en el 2014 también para México. Esto les permite comprar y vender las acciones de las cuatro plazas bursátiles, simplemente a través de un intermediario local (MILA, 2015).

Entre las ventajas más relevantes de este acuerdo, está el hecho de que ningún mercado pierde su independencia, ni autonomía regulatoria, pero se comprometen con el crecimiento en conjunto como mercado integrado, aprovechando las complementariedades de sus mercados (MILA, 2015).

Los países integrados en el MILA, en conjunto, conforman el primer mercado por número de emisores en América Latina, el segundo por capitalización bursátil y el tercero por volumen de negociación (BVC, 2015a).

#### **Colombia**

La Bolsa de Valores de Colombia S.A. es una empresa privada listada en el mercado de valores. La Bolsa administra plataformas de negociación de los mercados de Renta Variable, Renta Fija y Derivados Estandarizados, adicionalmente, a través de sus filiales, opera los mercados de divisas, commodities energéticos y servicios de proveeduría de precios (BVC, 2015b).

La Bolsa de Valores de Colombia tiene una capitalización de 262.101.261.149 US (cifra del 2014 a precios actuales) (Mundial, 2015).

El COLCAP, principal índice del mercado accionario colombiano (Banco de la República, 2015), es un indicador de capitalización bursátil que refleja los cambios de precios de las acciones con más liquidez dentro de la Bolsa de Valores de Colombia (BVC). Es esta, la participación de cada una de las acciones del índice se determina por el valor correspondiente de la capitalización bursátil ajustada (flotante de la compañía multiplicado por el último precio) (BVC, 2015c).

#### **México**

La Bolsa Mexicana de Valores tiene una capitalización de 525.056.676.421US (cifra del 2014 a precios actuales) (Mundial, 2015).

El índice mexicano IPC (Índice de Precios y Cotizaciones) es el indicador bursátil más importante de la Bolsa Mexicana de Valores. Es un índice de capitalización ponderada de las 35 acciones líderes que cotizan en la Bolsa Mexicana de Valores. El índice fue desarrollado con un nivel básico de 0,78 a partir del 30 de octubre 1978 (Bloomberg, 2015b).

En abril del 2012, la BMV modificó la metodología de selección del índice teniendo en cuenta las acciones con restricción de disponibilidad y estableciendo cinco filtros para las emisoras de la bolsa que conformen el indicador bursátil. Los filtros se establecieron en el siguiente orden: tiempo mínimo de operación continua; porcentaje de acciones flotantes mínimo; valor de capitalización flotado mínimo; mayor factor de rotación eligiendo sólo 45 emisoras; y finalmente, calificación conjunta sobre el factor de rotación, valor de capitalización flotado, mediana mensual del importe operado en la Bolsa Mexicana de Valores de los últimos 12 meses. De esta manera quedan seleccionadas las 35 emisoras del IPC siendo revisadas cada seis meses y sustituidas por otras de no cumplir con los requisitos de selección (ADVFN, 2015).

Otros indicadores bursátiles en la BMV, que utilizan criterios diferentes al IPC son el Índice de Mediana Capitalización (IMC30), formado por empresas con un nivel de capitalización alto pero no lo suficiente para hacer parte del IPC; el Inmex, que representa entre las 20 y 25 empresas con mayor capitalización y el Índice Habita (IH), formado por las empresas principales dentro del sector de la construcción y vivienda (Rankia, 2012).

### **Chile**

Los mercados de la Bolsa de Santiago actualmente transan más de 2.000 millones de USD diarios en acciones, instrumentos de renta fija, títulos extranjeros, valores de intermediación financiera y cuotas de fondos (Bolsa de Comercio de Santiago, 2015).

La Bolsa de Comercio de Santiago tiene una capitalización de 313.325.267.335 US (cifra del 2014 a precios actuales) (Mundial, 2015).

El Índice de Precio selectivo de Acciones (IPSA) es el principal índice bursátil de Chile, compuesto por las 40 acciones con el mayor volumen promedio de negociación anual en la Bolsa de Comercio de Santiago. Este índice ha sido calculado desde 1977 y se actualiza trimestralmente (Bloomberg, 2015a).

Otros índices importantes de la Bolsa de Comercio de Santiago son el Índice General de Precios de las Acciones (IGPA), índice de carácter patrimonial cuyo

objetivo es medir las variaciones de todas las acciones listadas en bolsa; el INTER-10, que refleja las variaciones de precios de las diez acciones del OPSA que cotizan ADR's en el exterior y el ADRIán, que mide el cambio de precios de todas las acciones chilenas cotizadas en mercados internacionales por la totalidad de los ADR's tranzados en el exterior y los sus efectos locales (Chile, 2015).

## **Perú**

La Bolsa de Valores de Lima S.A. es una sociedad cuyo objetivo es facilitar la negociación de valores inscritos, facilitando los mecanismos, servicios y sistemas adecuados para la intermediación de instrumentos derivados, valores de oferta pública e instrumentos que no sean objeto de emisión masiva (BVL, 2015a).

El índice más importante y referente del mercado peruano es el S&P/BVL Perú General, que refleja la tendencia promedio de las principales acciones cotizadas en Bolsa. A partir del 4 de Octubre del 2015 será un índice de capitalización del free float. Este índice será de retorno total, es decir asumirá la reinversión de los dividendos.

Otros índices importantes de la BVL son el S&P/BVL Peru Select, también un índice de capitalización pero con mayores requerimientos de liquidez y del tamaño de la capitalización del free float que el S&P/BVL Perú General; el S&P/BVL Lima 25, índice basado en la liquidez pues mide el desempeño de las 25 acciones con mayor negociación dentro de la Bolsa de Valores de Lima y el Índice de Buen Gobierno Corporativo (IBGC), el cual tiene como objeto reflejar el comportamiento de los precios de las acciones de aquellas empresas listadas que adoptan buenas prácticas de gobierno corporativo (BVL, 2015b).

La Bolsa de Valores de Lima tiene una capitalización de 96.850.058.685 US (cifra del 2014 a precios actuales) (Mundial, 2015).

### **1.3.2.2 Plataformas en línea**

#### **Penetración del internet**

En el mundo se realizaban en el 2014 alrededor de 5,7 billones de búsquedas en Google (Statistic Brain Research Institute, 2015).

A nivel mundial la penetración del internet es del 45% y el crecimiento desde el año 2000 es del 806% (Stats, 2015).



**Tabla 1: Uso del Internet**

	<b>Penetración del internet</b>	<b>Tiempo promedio dedicado por persona al internet</b>	<b>Número de internautas</b>	<b>Penetración esperada 2018</b>
<b>Colombia</b>	50%	4,3 hrs/día	26,936,343	65%
<b>Chile</b>	64%	1,5 hrs/día	10,000,000	74%
<b>México</b>	48%	5 hrs/día	59,000,000	65%
<b>Perú</b>	56%	2 hs, 40 min/día	10,785,000	64%

(Telesemana, 2015; US Media Consulting, 2014)

### Google

La participación de Google dentro de los motores de búsqueda a nivel global es del 89% seguida por Yahoo (3,4%) y Bing (3,3%) (Statista, 2015).

**Tabla 2: Participación en total motores de búsqueda a nivel global de computadores**

	<b>Participación en total motores de búsqueda a nivel global de computadores</b>
<b>Google Colombia</b>	1,47%
<b>Google Chile</b>	0,2%
<b>Google México</b>	1,2%
<b>Google Perú</b>	0,63%

(Market Share Statistics for Internet Technologies, 2015)

## Google Search

Google Search es el motor de búsqueda más utilizado en la Web (Alexa, 2015). Como parte de sus servicios la compañía Google lanzó una plataforma gratuita, Google Trends, que analiza un porcentaje de las búsquedas de Google Search para determinar la frecuencia de búsqueda de términos específicos en relación con el total de búsquedas que cumplen con los mismos parámetros de tiempo y localización (Google, 2015a). Esta información es cuantificada mediante el siguiente índice de interés de búsqueda (Xu Rui, 2015):

$$\text{Search Interest} = \frac{\# \text{ de búsquedas de un término específico}}{\# \text{ total de búsquedas}} \quad (1)$$

Luego este índice de búsquedas del término específico es normalizado en una escala relativa de valores entre 0 y 100 mediante la división de cada valor del índice de interés de búsquedas entre el mayor valor obtenido por este mismo índice (Google, 2015b). Google Trends provee la información histórica de búsquedas de manera semanal.

## Wikipedia

Wikipedia es autodefinida como “un esfuerzo colaborativo por crear una enciclopedia gratis, libre y accesible por todos” (Wikipedia, 2015).

Wikipedia fue lanzada el 15 de enero de 2001. La edición en español se inició el 20 de mayo del mismo año, al ampliarse el proyecto Wikipedia desde la primera versión monolingüe en inglés hacia otros idiomas (Wikipedia, 2015).

Actualmente Wikipedia en español recibe alrededor de 37 millones de visitas al día convirtiéndola en la segunda en número de visitas mensuales luego de Wikipedia en inglés (Zachte, 2015) y diariamente se crean más de 400 artículos (Wikipedia, 2015).

### 1.3.2.3 Minería de Datos

Es un proceso utilizado para el descubrimiento de patrones desconocidos en grandes volúmenes de datos. Utiliza los métodos de la inteligencia artificial, aprendizaje automático, estadística y sistemas de bases de datos (Maletic & Marcus, 2005).

El objetivo general del proceso de minería de datos consiste en extraer información de un conjunto de datos y transformarla en una estructura comprensible para su uso posterior ([Equihua, 2014](#)).

La minería de datos es un proceso englobado dentro de la tendencia de Big Data que, de acuerdo a IBM, es:

La tendencia en el avance de la tecnología que ha abierto las puertas hacia un nuevo enfoque de entendimiento y toma de decisiones, la cual es utilizada para describir enormes cantidades de datos (estructurados, no estructurados y semi estructurados) que tomaría demasiado tiempo y sería muy costoso cargarlos a una base de datos relacional para su análisis. De tal manera que, el concepto de Big Data aplica para toda aquella información que no puede ser procesada o analizada utilizando procesos o herramientas tradicionales (Fragoso, 2012).

#### **1.3.2.4 Estadística**

##### **Estadística descriptiva**

La estadística descriptiva es un conjunto de técnicas que buscan organizar, analizar y presentar un conjunto de datos o una muestra (Fernández, Cordero & Córdoba, 2002). Esta ofrece un sentido de la ubicación del centro de los datos, de su variabilidad y de la naturaleza general de la distribución de observaciones en una muestra (Walpole, 2012).

##### **Distribuciones de probabilidad**

Las variables aleatorias pueden definirse como cantidades que pueden tomar valores diferentes entre una y otra prueba de un experimento, el resultado exacto de un evento aleatorio. La distribución de probabilidad de una variable aleatoria enlista todos los posibles valores que esta puede tomar y la probabilidad de cada uno (Hanke & Reitsch, 1996).

##### **Pruebas de hipótesis**

Una hipótesis estadística es una afirmación acerca de la distribución de una o más variables aleatorias. Una prueba de hipótesis estadística es un procedimiento que busca determinar si se acepta una hipótesis, denominada nula, o se rechaza en favor de otra, conocida como alternativa (Freund, Miller & Miller, 2000).

Las pruebas de hipótesis se pueden resumir en los siguientes pasos:

#### Relación entre los patrones de búsqueda de internet y los movimientos del MILA

1. Formular las hipótesis nula y alternativa.
2. Seleccionar un estadístico de prueba con una distribución muestral apropiada.
3. Determinar una región de rechazo de tamaño  $\alpha$ .
4. Determinar el valor del estadístico de prueba a partir de los datos de la muestra.
5. Comprobar si el valor del estadístico de prueba cae en la región de rechazo y con base en ello aceptar o rechazar la hipótesis nula.

El valor-p es el menor nivel de significancia en el cual la hipótesis nula podría haberse rechazado, por lo tanto una hipótesis nula es rechazada cuando el valor p sea mayor que  $\alpha$  ( Freund, Miller & Miller , 2000).

Algunas pruebas de hipótesis son:

##### Kolmogorov-Smirnov

La prueba de Kolmogorov-Smirnov es una prueba no paramétrica que se utiliza para determinar la bondad de ajuste entre dos distribuciones, permite saber si los datos provienen de una población que tiene una distribución específica. Esta prueba es frecuentemente utilizada para determinar si los datos analizados siguen una distribución normal (Freund, Miller & Miller , 2000).

##### Wilcoxon Rank-sum

La prueba de Wilcoxon Rank-sum hace parte de las pruebas no paramétricas y es utilizada para comparar la mediana de dos muestras que están relacionadas y para determinar si hay diferencias entre ellas. Se utiliza cuando no se puede suponer la normalidad de las muestras, como una alternativa a la prueba t-student (Freund, Miller & Miller , 2000).

##### Kruskal-Wallis

La prueba de Kruskal-Wallis es una prueba no paramétrica que se utiliza para determinar si un grupo de datos proviene de la misma población (Freund, Miller & Miller , 2000).

#### 1.4 HIPÓTESIS

Existe una relación entre los patrones de búsqueda de palabras claves en Google y artículos en Wikipedia de algunas empresas listadas en bolsa y el

comportamiento del mercado accionario en los países perteneciente al Mercado Integrado Latinoamericano.

Algunas teorías psicológicas (Simon, 1957) muestran que los inversionistas tienden a presentar mayor aversión a perder dinero que a dejar de ganar dinero. Adicionalmente, se sabe que la búsqueda de información requiere de tiempo y esfuerzo. Por lo tanto, estudios sugieren que en momentos de incertidumbre financiera, las personas tienden a buscar más información sobre sus inversiones, reflejando el miedo a una pérdida de dinero (Moat, 2013). Este proceso búsqueda de información es seguido por una decisión de venta.

Estas búsquedas se llevan a cabo a través de plataformas como Google y Wikipedia, por lo que un aumento en el interés por una compañía específica prevee una futura caída en el precio de su acción.

## 2. METODOLOGÍA

Se utilizarán técnicas de minería de datos para convertir una gran cantidad de datos sobre los patrones de búsqueda de artículos de Wikipedia y palabras clave en Google en información relevante que pueda ser interpretada con mayor facilidad. Esto se realizará con la implementación de diferentes metodologías estadísticas en las que se incluye el análisis descriptivo de datos y conceptos de inferencia estadística.

Este proceso estará basado en la metodología de Preis, Moat y Stanley (2011) y será realizado siguiendo el procedimiento descrito a continuación:

Para llevar a cabo el primer objetivo específico:

Determinar los activos listados en bolsa, los artículos de Wikipedia y las palabras claves de búsqueda en Google que se emplearán para el estudio.

Se realizarán las siguientes actividades:

- A partir del índice representativo del mercado de cada país del MILA, elegir los activos que cumplan con los criterios establecidos, con el fin de usarlos como criterio de búsqueda.
- Establecer los artículos de Wikipedia que mejor describan la empresa asociada a cada activo seleccionado, teniendo en cuenta el contenido de los artículos.
- Definir las palabras clave de búsqueda de Google asociadas a cada activo.

Se definirá como “Palabra Clave” el Tema que mejor describa a la empresa emisora del activo, en caso de no existir un Tema, se elegirá el Término de Búsqueda que mejor describa al activo.

Para llevar a cabo el segundo objetivo específico:

Implementar una estrategia hipotética de inversión con la información obtenida.

Se realizarán las siguientes actividades:

- Obtener los datos para las visitas a los artículos de Wikipedia mediante el uso del paquete *Wikipediatrend* del programa estadístico R que provee acceso a la información histórica de Wikipedia en la página [stats.grok.se](http://stats.grok.se).

Relación entre los patrones de búsqueda de internet y los movimientos del MILA

- Obtener los datos para el número de búsquedas de las palabras clave en Google mediante la herramienta *Google Trends*.
- Obtener los precios históricos de cada activo seleccionado, para el periodo definido, a través de Bloomberg.
- Construir el programa en R para la implementación de la siguiente metodología introducida por Preis, Moat y Stanley (2013):

Los artículos y términos de búsqueda de las empresas están asociados a su respectiva acción, la estrategia se realizará de manera independiente para cada activo y los análisis de los resultados serán analizados por país.

Para cuantificar la actividad de los usuarios en Wikipedia y Google y siguiendo la metodología de Preis, Moat y Stanley (2013), se utilizará el promedio de las visitas y búsquedas  $n(t)$ , realizadas durante la semana ( $t$ ) y para cuantificar el cambio en los patrones de búsqueda se utilizará  $\Delta n$  donde:

$$\Delta n(t, \Delta t) = n(t) - N(t - 1, \Delta t) \quad (2)$$

$$N(t - 1, \Delta t) = \frac{n(t - 1) + n(t - 2) + \dots + n(t - \Delta t)}{\Delta t} \quad (3)$$

La estrategia se implementará con el precio de cierre  $P(t)$  del primer día de negociación de la semana  $t$ . De forma que, si disminuye el volumen de búsquedas tal que,

$$\Delta n(t, \Delta t) \leq 0 \quad (4)$$

Se tomará una posición larga comprando el índice a un precio  $P(t + 1)$  y vendiéndolo a un precio  $P(t + 2)$ .

Luego se calculará el retorno continuo  $R(t)$  de la estrategia, donde

$$R(t) = \text{Log}[P(t + 2)] - \text{Log}[P(t + 1)] \quad (5)$$

Por otro lado, si el volumen de búsquedas aumenta tal que:

$$\Delta n(t, \Delta t) > 0 \quad (6)$$

#### Relación entre los patrones de búsqueda de internet y los movimientos del MILA

Se tomará una posición corta vendiendo el índice a un precio  $P(t + 1)$  y cerrando la posición al comprar a un precio  $P(t + 2)$ .

Luego se calculará el retorno continuo  $R(t)$  de la estrategia, donde:

$$R(t) = \text{Log}[P(t + 1)] - \text{Log}[P(t + 2)] \quad (7)$$

Adicionalmente se implementará una estrategia de trading aleatoria que simulará 10.000 iteraciones independientes para el mismo periodo de estudio, donde la probabilidad de tomar una posición corta o larga será siempre del 50%.

Para la determinación del  $\Delta t$  se considerarán periodos entre 1 y 6 semanas aplicando la metodología descrita anteriormente para cada uno de ellos. Se tomará el periodo que tenga el mayor retorno promedio para utilizarlo en el resto del estudio.

Para llevar a cabo el tercer objetivo específico:

Validar estadísticamente la metodología aplicada.

Se realizarán las siguientes actividades:

- Identificar el tipo de distribución que siguen los retornos.

Para determinar si los retornos siguen una distribución normal, realizar la prueba de hipótesis Kolmogorov Smirnov si el tamaño de la muestra es mayor a 30 o la prueba de hipótesis Shapiro-Wilks si es menor a 30.

La prueba de normalidad utilizada tendrá la siguiente hipótesis nula y alternativa:

- Los retornos siguen una distribución normal.
- Los retornos no siguen una distribución normal.

Si los retornos no siguen una distribución normal crear la distribución de probabilidad para los retornos.

- Utilizar pruebas de hipótesis para determinar la significancia estadística de los retornos de la estrategia aleatoria de inversión.



#### Relación entre los patrones de búsqueda de internet y los movimientos del MILA

En caso de que los retornos sigan una distribución normal se utilizará una prueba t-student simple, de lo contrario se utilizará un test del signo de rangos de Wilcoxon.

La prueba a realizar tendrá la siguiente hipótesis nula y alternativa:

- La media de los retornos es estadísticamente igual a cero
- La media de los retornos es estadísticamente mayor a cero
- Determinar si existe una diferencia significativa entre la estrategia hipotética de inversión y la estrategia aleatoria.

En caso de que los retornos sigan una distribución normal se utilizará una prueba t-student simple, de lo contrario se utilizará la prueba del signo de rangos de Wilcoxon.

La prueba a realizar tendrá la siguiente hipótesis nula y alternativa:

- La diferencia entre la media de la estrategia hipotética de inversión y la media de la estrategia aleatoria es igual a cero.
- La diferencia entre la media de la estrategia hipotética de inversión y la media de la estrategia aleatoria es diferente de cero.

En caso de que se rechace la hipótesis nula, la diferencia entre la estrategia hipotética de inversión y la estrategia aleatoria estará determinada por la diferencia entre sus medias.

- Determinar las diferencias entre los retornos de los países.

En caso de que los retornos sigan una distribución normal se utilizará una prueba ANOVA de un factor, de lo contrario se utilizará el test de Kruskal-Wallis.

Se realizarán pruebas de hipótesis con las siguientes hipótesis nulas y alternativas:

- La diferencia entre la media de los retornos del país  $X$  y la media de los retornos del país  $Y$  es igual a cero.
- La diferencia entre la media de los retornos del país  $X$  y la media de los retornos del país  $Y$  es diferente de cero.



### 3. DESARROLLO DEL PROYECTO

#### 3.1 OBJETIVO 1

Seleccionar los activos listados en bolsa, los artículos de Wikipedia y las palabras claves de búsqueda en *Google* que se emplearán para el estudio.

##### 3.1.1 Determinar el índice representativo del mercado de cada país del MILA

Se tomó como referencia la bolsa de valores de cada uno de los países del MILA para determinar cuál era el índice representativo de cada mercado. Los resultados fueron los siguientes:

País	Índice Representativo
Colombia	COLCAP
Chile	IPSA
Perú	S&P/BVL Perú General
México	MexBol

Tabla 3: Índices representativos por país

##### 3.1.2 Definir los activos que se utilizarán

Se tomaron todos los activos dentro del índice representativo de cada país que cumplieras con los siguientes criterios:

- El activo debe tener, por lo menos, un artículo en Wikipedia (en inglés o en español). De lo contrario no se tomará en cuenta.
- Los datos de tendencia de búsqueda en *Google Trends* deben ser semanales. Nota: cuando el término de búsqueda no es muy popular los datos aparecen únicamente mensuales.
- Si dentro de los activos de cada índice se encuentra tanto la acción ordinaria como la preferencial de una misma empresa se elegirá únicamente la acción ordinaria. Por ejemplo, dentro del índice COLCAP se

#### Relación entre los patrones de búsqueda de internet y los movimientos del MILA

encuentran los activos “BCOLO” y “PFBCOLO”. En este caso se elegirá “BCOLO” y se eliminará “PFBCOLO”.

- Si dentro de los activos de cada índice se encuentra más de una acción que represente a la misma empresa o grupo empresarial, se eliminará aquel activo con menor ponderación dentro del índice. Por ejemplo, dentro del índice COLCAP se encuentran los activos “GRUPOARG” y “CEMARGOS” con ponderaciones dentro del índice de 6,3% y 5,1% respectivamente. Se eliminará de la selección “CEMARGOS” por tener un menor peso en el COLCAP.

#### Definición activos índice COLCAP

Según los criterios establecidos se eliminaron los siguientes activos del índice:

- PFBCOLO, PFGRUPSU, PFGRUPOA y PFCEMARG debido a la existencia de sus respectivas acciones ordinaria dentro del índice.
- CEMARGOS debido a la existencia de una acción que representa a la misma empresa o grupo empresarial y que tiene mayor ponderación en el índice (GRUPOARG).
- EEB, ISA y CNEC debido a que no tienen artículo en Wikipedia.

Los activos seleccionados de este índice fueron los siguientes:

GRUPOSUA	PFAVAL	EXITO	PFAVH
ECOPETL	BCOLO	BOGOTA	BVC
NUTRESA	CORFICOL	PFDVVND	PREC

Relación entre los patrones de búsqueda de internet y los movimientos del MILA

GRUPOARG	ISAGEN	CLH	
----------	--------	-----	--

Tabla 4: Activos Colombia

**Definición de activos índice MexBol**

- PINFRA\*, SITESL, ICHB y SIMECB debido a que no tienen artículo en Wikipedia.
- ASURB y OMAB debido a la existencia de una acción que representa a la misma empresa o grupo empresarial y que tiene mayor ponderación en el índice (GAPB).

Los activos seleccionados de este índice fueron los siguientes:

FEMSAUBD	CEMEXCPO	BIMBOA	ELEKTRA*	PE&OLES*
AMXL	ALFAA	GRUMAB	AC*	IENOVA*
GFNORTEO	GAPB	GFINBURO	ALSEA*	OHLMEX*
TLEVICPO	KOFL	LIVEPOLC	LALAB	LABB
WALMEX*	KIMBERA	MEXCHEM*	GCARSOA1	GFREGIO
GMEXICOB	SANMEXB	GENTERA*	ICA*	

Tabla 5: Activos México

**Definición de activos índice IPSA:**

Según los criterios establecidos se eliminaron los siguientes activos del índice:

1. SECUR y BESALCO debido a que no tienen artículo en Wikipedia.
2. CCU, ECL, QUINENC, BANMED, SMSAAM, ILC, CAP, FORUS, BUPACL y EMBONOB debido a la carencia de datos semanales en *Google Trends*.

**Relación entre los patrones de búsqueda de internet y los movimientos del MILA**

Los activos seleccionados de este índice fueron los siguientes:

ENERSIS	LAN	SMCHILEB	CONCHA
COPEC	COLBUN	AESGENER	IAM
ENDESA	CHILE	ANTAR	GASCO
FALAB	BCI	ENTEL	RIPLEY
BSAN	SQM/B	PARAUCO	SK
CMPC	CORPBANC	SONDA	VAPORES
CENCOSUD	AGUAS/A	ANDINAB	SALFACOR

**Tabla 6: Activos de Chile**

**Definición de activos índice S&P/BVL Perú General:**

- CORAREI1 debido a la existencia debido a la existencia de una acción que representa a la misma empresa o grupo empresarial y que tiene mayor ponderación en el índice (CORAREC1).
- AIHC1, EXALMC1, UNACEMC1, ENERSUC1, EDELNOC1 y LGC debido a que no tienen artículo en Wikipedia.
- MIRL, POMALCC1, EDEGELC1, VOLCABC1, PML, ATACOBC1, CASAGRC1, GRAMONC1, FERREYC1, INRETC1, INVCENC1, TV, BAP, RELAPAC1, AUSTRAC1 y BROCALC1 debido a la carencia de datos semanales en *Google Trends*.

## Relación entre los patrones de búsqueda de internet y los movimientos del MILA

Los activos seleccionados de este índice fueron los siguientes:

CORAREC1	BACKUSI1	MINSURI1
MILPOC1	IFS	LUSURC1
SCCO	CVERDEC1	BVN
ALICORC1	CPACASC1	

**Tabla 7: Activos Perú**

### 3.1.3 Establecer los artículos de Wikipedia para cada activo seleccionado

Para la selección de los artículos se utilizaron los siguientes criterios:

- Se tomó el artículo de Wikipedia cuyo título fuera igual al nombre de la empresa emisora de la acción.
- En caso de haber un artículo en inglés y otro en español, se tomaron ambos.

Los artículos que fueron elegidos para cada activo se encuentran en el Anexo 1.

### 3.1.3 Definir las palabras clave asociadas a cada activo

Las palabras clave en *Google Trends* pueden ser de dos tipos:

- Términos de Búsqueda: son aquellas que solo tienen en cuenta las búsquedas que incluyen esa cadena de texto específica
- Temas: en estas, el algoritmo de *Google Trends* tiene en cuenta diferentes búsquedas relacionadas con el mismo tema.

Para cada activo se eligió Tema relacionado con la empresa emisora y en caso de no tener Tema, se eligió el Término de Búsqueda que mejor representara a la empresa.

Los temas y términos de búsqueda que fueron elegidas para cada activo se encuentran en el Anexo 1.

### 3.2 OBJETIVO 2

Implementar una estrategia hipotética de inversión con la información obtenida previamente.

#### 3.2.1 Obtener los datos de los activos

Según lo presentado anteriormente, se obtuvieron datos para 15 activos en Colombia, 29 en México, 27 en Chile y 11 en Perú.

Los datos semanales de las tendencias de búsqueda en *Google* fueron descargados de *Google Trends* ([www.google.com/trends/](http://www.google.com/trends/)) y los datos diarios de visitas de los artículos de Wikipedia fueron descargados de *Wikipedia Article Traffic Statistics* ([stats.grok.se](http://stats.grok.se)).

Los datos de ambas fuentes se obtuvieron para un periodo que comienza el 5/01/2008 y termina el 2/01/2016.

Para este periodo, los precios de cierre de las acciones el día lunes de cada semana se descargaron de la plataforma *Bloomberg*.

#### 3.2.2 Desarrollar el código en R para la implementación de la metodología

Debido a que los datos de búsquedas en *Google* solo estaban disponibles de forma semanal y aquellos en *Wikipedia* de forma diaria, se calculó el promedio semanal para los datos de *Wikipedia*, teniendo en cuenta que la fecha de inicio y de final de semana fueran iguales.

Las semanas se eligieron empezando el domingo y terminando el sábado pues así las presentaban los datos de *Google Trends*.

Inicialmente se analizaron los precios de cierre  $P(t)$  de cada activo en el primer día de negociación de la semana  $t$ , con el fin de determinar la relación entre el volumen de búsquedas de un Término de Búsqueda o Tema y la dirección de las decisiones de negociación.



#### Relación entre los patrones de búsqueda de internet y los movimientos del MILA

Se utilizó *Google Trends* para determinar cuántas búsquedas  $n(t-1)$  se realizaron para un término específico en la semana  $t-1$ .

Para estudiar si existía una correlación entre las tendencias de búsquedas en *Google*, capturadas por *Google Trends*, y las visitas a los artículos en Wikipedia con los cambios futuros en el precio de los activos entre el 2008 y el 2016, se decidió implementar una estrategia hipotética de inversión utilizando estos datos.

Se tomó en cuenta que sólo se pueden obtener ganancias en una estrategia de inversión si los cambios futuros en el precio de la acción son anticipados correctamente.

Además, se estableció una estrategia de inversión teniendo en cuenta algunas teorías psicológicas (Simon, 1957) que muestran que ante momentos de incertidumbre financiera las personas tienden a buscar más información sobre sus inversiones y que esta búsqueda es seguida de una decisión de venta.

La estrategia hipotética de inversión se implementa vendiendo la acción al precio de cierre  $P(t)$  en el primer día de la semana ( $t$ ), si  $\Delta n(t, \Delta t) > 0$ , y comprando la acción al precio  $P(t+1)$  al final del primer día de negociación de la semana siguiente.

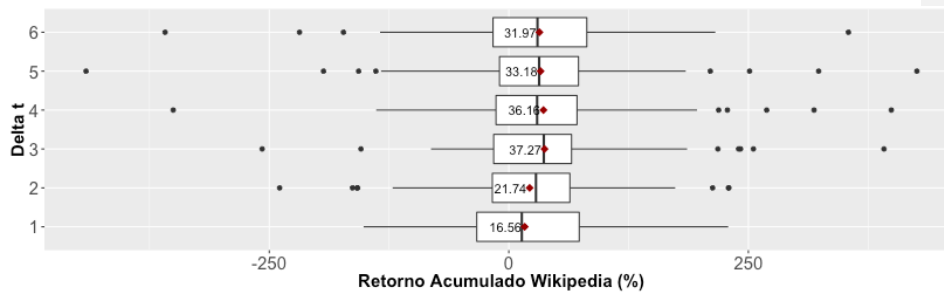
Si  $\Delta n(t, \Delta t) \leq 0$ , se compra la acción al precio de cierre  $P(t)$  en el primer día de la semana ( $t$ ), y se vende la acción al precio  $P(t+1)$  al final del primer día de negociación de la semana siguiente.

Al realizar la negociación bajo esta estrategia se realizan aproximadamente 104 transacciones anuales, una compra y una venta semanal. Para este estudio se omiten los costos de transacción, sin embargo, es importante tener en cuenta que, en la implementación real de la estrategia, estos costos tendrían un impacto en los retornos.

Si se toma una posición corta, vendiendo al precio de cierre  $P(t)$  y comprando al precio  $P(t+1)$  el retorno acumulado  $R$  cambia en  $\log(P(t)) - \log(P(t+1))$ . Si se toma una posición larga, comprando al precio de cierre  $P(t)$  y vendiendo al precio  $P(t+1)$ , el retorno acumulado  $R$  cambia en  $\log(P(t+1)) - \log(P(t))$ .

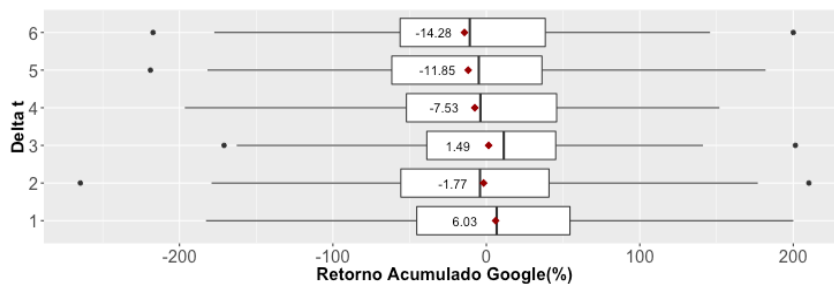
Se realizó un análisis previo de los retornos para determinar el  $\Delta t$  que maximizan el rendimiento. Los resultados fueron los siguientes:

# Relación entre los patrones de búsqueda de internet y los movimientos del MILA



**Imagen 1:  $\Delta t$  Wikipedia**

Este diagrama de caja muestra los diferentes cuartiles de los retornos acumulados de los activos para cada  $\Delta t$  desde  $t=1$  semanas hasta  $t=6$  semanas, basado en las visitas de Wikipedia. Los retornos fueron calculados de forma logarítmica. El punto rojo representa el promedio de los retornos acumulados para cada  $\Delta t$  y la línea vertical muestra la mediana. El  $\Delta t$  con mayor retorno es  $\Delta t=3$  semanas con un retorno acumulado promedio de 37.27%.

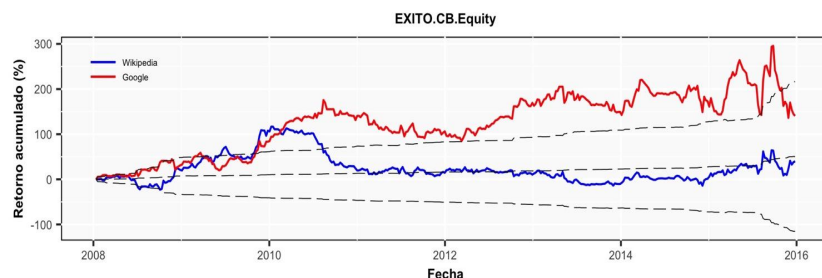


**Imagen 2:  $\Delta t$  Google**

Este diagrama de caja muestra los diferentes cuartiles de los retornos acumulados de los activos para cada  $\Delta t$  desde  $t=1$  semanas hasta  $t=6$  semanas, basado en las visitas de Google. Los retornos fueron calculados de forma logarítmica. El punto rojo representa el promedio de los retornos acumulados para cada  $\Delta t$  y la línea vertical muestra la mediana. El  $\Delta t$  con mayor retorno es  $\Delta t=1$  semanas con un retorno acumulado promedio de 6.03%.

## Relación entre los patrones de búsqueda de internet y los movimientos del MILA

Luego de determinar los  $\Delta t$  para los datos de Wikipedia y *Google* que maximizan el retorno acumulado promedio se implementó, para cada uno de los activos, una estrategia hipotética de inversión y se comparó con una estrategia aleatoria en la cual la probabilidad de tomar una posición larga o corta era igual (ambas del 50%) e independiente de las decisiones anteriores. Para obtener la media y la desviación estándar de la estrategia aleatoria se realizaron 10.000 simulaciones mediante el método de Monte Carlo. Los resultados para cada activo se muestran en el ANEXO 2.



**Imagen 3: Retornos acumulados Éxito**

En la gráfica se muestran los retornos acumulados obtenidos aplicando la estrategia hipotética de inversión para los datos de Wikipedia, con un  $\Delta t=3$  semanas y para los datos de Google, con un  $\Delta t=1$  semana. Adicionalmente se presenta la media y la desviación estándar de la estrategia aleatoria.

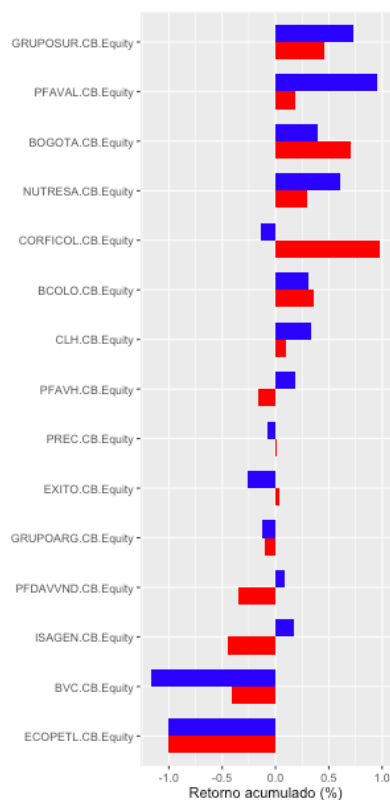
La imagen 3 refleja que la estrategia hipotética de inversión propuesta, aplicada al activo EXITO, con un  $\Delta t=3$  semanas para Wikipedia y un  $\Delta t=1$  semana para Google, habría aumentado el valor de la inversión en un 40% y 140% respectivamente.

Para darle más validez a los datos, se calculó el rendimiento de cada estrategia hipotética de inversión utilizando la media de los seis retornos obtenidos para  $\Delta t=1$  hasta  $\Delta t=6$ . Además, los retornos para cada estrategia hipotética se calcularon como el logaritmo de los cambios relativos de precios para cada activo.

En las imágenes 4, 5, 6 y 7 se muestran los retornos acumulados para cada estrategia, en términos de la desviación estándar. Las barras rojas representan la estrategia hipotética de inversión realizada a partir de los datos de búsquedas en

#### Relación entre los patrones de búsqueda de internet y los movimientos del MILA

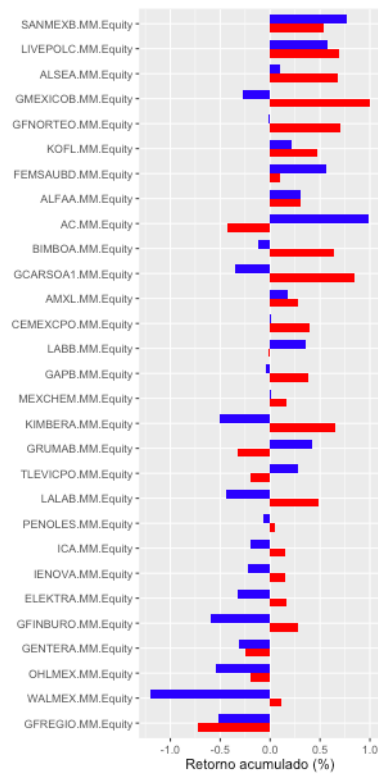
Google y las barras azules muestran los retornos obtenidos al implementar la estrategia hipotética basada en las búsquedas de artículos de *Wikipedia*.



**Imagen 4: Retornos acumulados Colombia**

Se representan los retornos promedio acumulados en términos de la desviación estándar de la estrategia aleatoria para cada activo de Colombia, siguiendo la estrategia hipotética de inversión con  $\Delta t=1$  semana a  $\Delta t=6$  semanas.

## Relación entre los patrones de búsqueda de internet y los movimientos del MILA



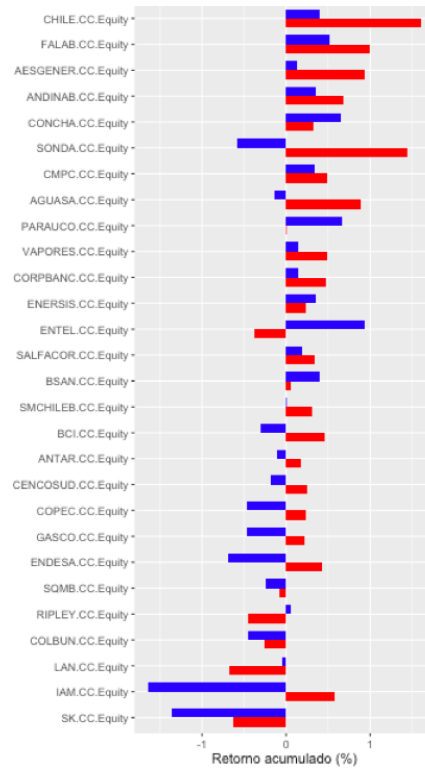
**Imagen 5: Retornos acumulados México**

Se representan los retornos promedio acumulados en términos de la desviación estándar de la estrategia aleatoria para cada activo de México, siguiendo la estrategia hipotética de inversión con  $\Delta t=1$  semana a  $\Delta t=6$  semanas.

M 10/11/16 6:31 PM

Formatted: Font color: Auto

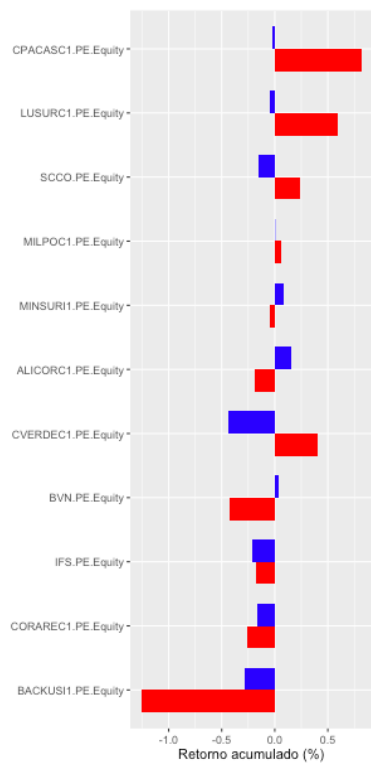
# Relación entre los patrones de búsqueda de internet y los movimientos del MILA



**Imagen 6: Retornos acumulados Chile**

Se representan los retornos promedio acumulados en términos de la desviación estándar de la estrategia aleatoria para cada activo de Chile, siguiendo la estrategia hipotética de inversión con  $\Delta t=1$  semana a  $\Delta t=6$  semanas.

# Relación entre los patrones de búsqueda de internet y los movimientos del MILA



**Imagen 7: Retornos acumulados Perú**

Se representan los retornos promedio acumulados en términos de la desviación estándar de la estrategia aleatoria para cada activo de Perú, siguiendo la estrategia hipotética de inversión con  $\Delta t=1$  semana a  $\Delta t=6$  semanas.

En la Tabla 8 se muestran los retornos acumulados promedio para  $\Delta t=1$  hasta  $\Delta t=6$  agrupados por país, para la estrategia hipotética de Google y Wikipedia.

Relación entre los patrones de búsqueda de internet y los movimientos del MILA

Estrategia	País	Retorno Acumulado
Google	Colombia	1.9%
	México	-3.6%
	Chile	-5.2%
	Perú	-18.5%
Wikipedia	Colombia	3.3%
	México	51.7%
	Chile	46.3%
	Perú	-6.1%

Tabla 8: Retornos Acumulados por Estrategia y País

### 3.3 OBJETIVO 3

Validar estadísticamente los resultados de la estrategia implementada anteriormente.

#### 3.3.1 El resultado de las pruebas realizadas con su respectivo análisis.

Para validar estadísticamente los resultados obtenidos, se realizó una prueba *T-Student* para los retornos acumulados de cada país y para cada estrategia de inversión. Estas pruebas tienen las siguientes características:

1. Se utilizó el promedio de los retornos acumulados logarítmicos para  $\Delta t=1$  hasta  $\Delta t=6$ , en términos de la desviación estándar de cada uno.
2. Se tomó un nivel de confianza del 95%
3. Se determinaron las siguientes hipótesis:
  - a. Hipótesis nula: la media de los retornos es igual de cero.
  - b. Hipótesis alternativa: la media de los retornos es mayor a cero.

Se obtuvieron los siguientes resultados:



Relación entre los patrones de búsqueda de internet y los movimientos del MILA

Estrategia	País	Valor-P	Resultado
Google	Colombia	0.4639	La media de los retornos no es estadísticamente mayor a cero
	México	0.5952	La media de los retornos no es estadísticamente mayor a cero
	Chile	0.634	La media de los retornos no es estadísticamente mayor a cero
	Perú	0.8902	La media de los retornos no es estadísticamente mayor a cero
Wikipedia	Colombia	0.4285	La media de los retornos no es estadísticamente mayor a cero
	México	<b>0.0007612</b>	La media de los retornos es estadísticamente mayor a cero
	Chile	<b>0.0029</b>	La media de los retornos es estadísticamente mayor a cero
	Perú	0.5689	La media de los retornos no es estadísticamente mayor a cero

Tabla 9: Resultados de Pruebas Estadísticas

### 3.3.2 Conclusiones generales sobre las relaciones entre los patrones de búsqueda de internet y los movimientos en el mercado seleccionado.

En la tabla 8 se puede observar que, los resultados de los retornos acumulados de *Google* fueron positivos únicamente para los activos de Colombia y los resultados de los retornos acumulados de *Wikipedia* fueron positivos para los activos de México, Colombia y Chile.

Sin embargo, luego de realizar las pruebas de hipótesis cuyos resultados se muestran en la Tabla 9, se puede concluir que los únicos retornos acumulados estadísticamente mayores a cero fueron aquellos obtenidos para México y Chile utilizando la estrategia hipotética de inversión con datos de visitas a los artículos de Wikipedia.

#### Relación entre los patrones de búsqueda de internet y los movimientos del MILA

Por lo tanto no hay una evidencia contundente de que la relación propuesta en la hipótesis de este estudio exista, es decir, la teoría de que ante momentos de incertidumbre financiera las personas tienden a buscar más información sobre sus inversiones y que esta búsqueda es seguida de una decisión de venta Moat (2013), no es válida para los países estudiados.

Otros estudios realizados para el mercado de Estados Unidos detectaron que el incremento de volumen de búsquedas en *Google Trends* (Preis, Moat, & Stanley, 2013) y *Wikipedia* (Moat et al., 2013) es seguido por caídas en el mercado accionario. Para trabajos futuros, sería interesante evaluar si la diferencia en los resultados entre Estados Unidos y los países pertenecientes al Mercado Integrado Latinoamericano podría ser explicada por los siguientes factores:

1. El mercado accionario de Estados Unidos en términos de la capitalización bursátil es aproximadamente 24 veces más grande que el mercado accionario de los países del MILA (Desjardins, 2016) y la penetración del internet en Estados Unidos es del 88.5 % mientras que aquella en México, Colombia, Chile y Perú es de 45.1 %, 56.9 %, 77.8 % y 41 % respectivamente (Internet Live Stats, 2016). Lo anterior implica un menor número de datos para este estudio en comparación con los estudios realizados para Estados Unidos.
2. La desigualdad social en Estados Unidos es menor que aquella en los países estudiados (World Bank, 2013) y su poder adquisitivo es mayor, por lo que en Estados Unidos una mayor parte de la población pertenece a la clase media y posiblemente tiene recursos para invertir pero no los suficientes para pagar un asesor financiero, utilizando el internet como fuente de información.
3. Es importante tener en cuenta que *Google* y *Wikipedia* no son las únicas fuentes de información utilizadas a la hora de tomar decisiones de inversión. Las personas tienen otras opciones como buscar en *Bloomberg*, consultar con un asesor financiero o invertir a través de fondos de inversión.

#### 4. DISCUSIÓN DE LOS RESULTADOS

El estudio tenía como objetivo principal determinar si existe la relación propuesta en la hipótesis entre las búsquedas en *Google* y *Wikipedia* y los movimientos del mercado accionario en los países pertenecientes al MILA.

Inicialmente se pretendían utilizar todos los activos que hicieran parte del índice representativo de cada uno de los países seleccionados para el estudio, con el objetivo de tener la mayor cantidad de datos posible para el análisis. Pero a medida que se fueron buscando los datos requeridos se vio la necesidad de establecer ciertas condiciones y filtros para la determinación de los activos a utilizar como por ejemplo la disponibilidad de datos en *Google Trends* o la existencia de un artículo representativo en *Wikipedia*. Esto redujo la base de datos para el estudio pero permitió utilizar datos uniformes y relevantes. Se seleccionaron 15 activos para Colombia, 29 para México, 27 para Chile y 11 para Perú.

Luego se implementó una estrategia hipotética de inversión siguiendo la teoría psicológica de que ante momentos de incertidumbre financiera las personas tienden a buscar más información sobre sus inversiones y que esta búsqueda es seguida de una decisión de venta. En esta estrategia de inversión si el número de visitas aumentaba en la semana  $t$ , se tomaba una posición corta vendiendo el activo en la semana  $t + 1$  y luego se cerraba esta posición comprándolo en la semana  $t + 2$ . Mientras que cuando el número de visitas bajaba, se tomaba una posición larga comprando el activo en la semana  $t + 1$  y luego se vendía en la semana  $t + 2$ . Esta estrategia de implemento utilizando los precios históricos de cierre de los activos seleccionados desde 5/01/2008 hasta 2/01/2016 y a partir de ella se calculó el retorno acumulado para cada activo y se calculó el retorno acumulado por país. Se obtuvieron resultados mixtos: a partir de los datos de *Google Trends* el único retorno positivo fue aquel de Colombia y a partir de los datos de *Wikipedia* los retornos positivos fueron aquellos de México y Chile.

Para validar estadísticamente estos resultados se realizó una prueba *T-Student* la cual determinó que los únicos retornos estadísticamente mayores a cero fueron aquellos calculados a partir de los datos de *Wikipedia* para Chile y México.

## 5. CONCLUSIONES Y CONSIDERACIONES FINALES

A partir de los resultados obtenidos para cada objetivo específico, se concluye que no se puede afirmar que existe la relación presentada en la hipótesis entre las búsquedas en *Google* y *Wikipedia* y los movimientos del mercado accionario de los activos pertenecientes al MILA.

Existen estudios realizados para el mercado de Estados Unidos que demuestran que la relación estudiada en esta investigación, es decir que incrementos en el volumen de búsqueda son seguidos por caídas en el mercado accionario, es cierta (Preis, Moat, & Stanley, 2013) pero también se han realizado estudios para otros mercados financieros como la Bolsa Alemana Xetra, que demuestran la relación contraria entre el volumen de búsquedas y los movimientos del mercado (Bank, Larch, & Peter, 2011). Por lo tanto sería interesante explorar en futuras investigaciones una relación inversa entre el volumen de búsquedas y los movimientos en el mercado accionario de México, Colombia, Chile y Perú.

Además se sugiere realizar este mismo estudio dentro de algunos años, cuando la penetración del internet incremente en los países estudiados y cuando el mercado accionario de estos países sea más maduro, eliminando posibles fallas presentadas en esta investigación como el limitado número de datos para los países del Mercado Integrado Latinoamericano comparado con el número de datos utilizado en la investigación aplicada a Estados Unidos (Preis, Moat, & Stanley, 2013).

Igualmente esta investigación podría ser realizada utilizando datos de otras fuentes como *Bloomberg*, plataforma que podría resultar más representativa como fuente de búsqueda a la hora de invertir.

En el presente estudio se exploró una posible relación utilizando una estrategia de inversión definida, propuesta en la literatura y aplicada en el pasado para Estados Unidos. Sin embargo, existen múltiples estrategias de inversión que podrían implementarse en futuras investigaciones para encontrar relaciones entre los patrones de búsqueda de internet y los movimientos en el mercado accionario.

Finalmente, para futuros estudios, sería interesante aplicar la estrategia hipotética de inversión utilizada en este trabajo en mercados financieros con un mayor volumen de acciones transadas y con una mayor penetración del internet (mas búsquedas en *Google* y *Wikipedia*) que los países del MILA, por ejemplo Canadá o Inglaterra.

En el transcurso del estudio se presentaron algunas limitaciones, principalmente con respecto al número de datos para el mercado de los países estudiados y a la disponibilidad de estos.

Aun cuando esta estrategia resultó exitosa para el mercado de Estados Unidos (Preis, Moat, & Stanley, 2013), los resultados de este estudio muestran que no lo es para los países del Mercado Integrado Latinoamericano. Esto comprueba la importancia de la validación numérica de una estrategia de inversión antes de aceptar como ciertos supuestos que pueden no ser ciertos en diferentes contextos.

## 6. **MATERIAL COMPLEMENTARIO**

Con el objetivo de presentar los resultados de este estudio se creó una aplicación en línea para exponer la información de forma gráfica, detallada y didáctica al usuario. Esta aplicación también permite seleccionar diferentes criterios como un país o un activo específico y visualizar los resultados correspondientes a las características seleccionadas.

Link de la aplicación: <https://tesis.shinyapps.io/app/>

## BIBLIOGRAFÍA

- Alexa. (2015). How popular is Google?. Retrieved from  
:http://www.alexa.com/siteinfo/google.com+yahoo.com+altavista.com
- ADVFN. (2015). ADVFN. Retrieved from http://mx.advn.com/ipc
- Alanyali, M., Moat, H. S., & Preis, T. (2013). Quantifying the relationship between financial news and the stock market. *Scientific Reports*, 3, 3578. doi:10.1038/srep03578
- Banco de la República. (2015). Banco de la República. Retrieved from  
http://www.banrep.gov.co/es/igbc
- Bank, M., Larch, M., & Peter, G. (2011). Search volume and its influence on liquidity and returns of German stocks.pdf. *Financ Mark Portfolio Manage*, 25(3), 239–264.
- Bloomberg. (2015a). Bloomberg. Retrieved from  
www.bloomberg.com/quote/IPSA:IND
- Bloomberg. (2015b). Bloomberg Business. Retrieved from  
www.bloomberg.com/quote/MEXBOL:IND
- Bolsa de Comercio de Santiago. (2015). Bolsa de Comercio de Santiago. Retrieved from  
http://www.bolsadesantiago.com/labolsa/Paginas/Qui%C3%A9nes-somos.aspx
- Bordino, I., Battiston, S., Caldarelli, G., Cristelli, M., Ukkonen, A., & Weber, I. (2012). Web search queries can predict stock market volumes. *PLoS ONE*, 7(7). doi:10.1371/journal.pone.0040014
- BVC. (2015a). Bolsa de Valores de Colombia. Retrieved from  
https://www.bvc.com.co/pps/tibco/portalbvc/Home/Mercados/descripciongeneral/acciones?action=dummy
- BVC. (2015b). Bolsa de Valores de Colombia. Retrieved from  
https://www.bvc.com.co/pps/tibco/portalbvc/Home/AcercaBVC/Perfil?action=dummy
- BVC. (2015c). Metodología para el cálculo del índice COLCAP. Retrieved from  
http://www.bvc.com.co/pps/tibco/portalbvc/Home/Mercados/descripciongeneral/indicesbursatiles?com.tibco.ps.pagesvc.action=updateRenderState&rp.currentDocumentID=13c20e3d\_13d5f9d729a\_7fcc0a0a600b&rp.revisionNumber=1&rp.attachmentPropertyName=Attachment&com.

- BVL. (2015a). Bolsa de Valores de Lima. Retrieved from  
web.archive.org/web/20110523124104/http://www.bvl.com.pe/acercalaempres  
a.html
- BVL. (2015b). Bolsa de Valores de Lima. Retrieved from  
www.bvl.com.pe/estadist/mercindicesmercado.html
- Chile, S. de V. y S. de. (2015). SVS educa. Retrieved from  
www.svs.cl/educa/600/w3-article-1589.html
- Choi, H., & Varian, H. (2012). Predicting the Present with Google Trends.  
*Economic Record*, 88(SUPPL.1), 2–9. doi:10.1111/j.1475-4932.2012.00809.x
- Desjardins, J. (2016, February 17). *All of the World's Stock Exchanges by Size*.  
Retrieved 09 20, 2016, from Visual Capitalist:  
http://www.visualcapitalist.com/all-of-the-worlds-stock-exchanges-by-size/
- Devore, J. L. (2008). Probabilidad y estadística para ingeniería y ciencias (Séptima  
Ed). CENGAGE Learning. Retrieved from  
www.uantof.cl/facultades/csbasicas/Matematicas/academicos/jreyes/DOCEN  
CIA/APUNTES/APUNTES PDF/Probabilidad y Estadística para Ingeniería y  
Ciencias - Jay Devore - Septima Edicion.pdf
- Equihua, S. (2014, January 14). *Data and Text Mining*. Retrieved October 11,  
2016, from Infotecarios: <http://www.infotecarios.com/data-text-mining/>
- Fernández, S., Cordero, J. M., & Córdoba, A. (2002). Estadística Descriptiva .  
Madrid: ESIC Ediciones.
- Fragoso, R. B. (2012). IBM developerWorks. Retrieved from  
https://www.ibm.com/developerworks/ssa/local/im/que-es-big-data/
- Francesco, D. a., & Marcucci, J. (2009). "Google it!" Forecasting the US  
unemployment rate with a Google job search index. *ISER Working Paper  
Series*. Retrieved from <http://econpapers.repec.org/paper/eseiserwp/2009-32.htm>
- Freund, J. E., Miller, I., & Miller, M. (2000) Estadística Matemática con  
Aplicaciones. 6ta Edición, México: Editorial Prentice Hall.
- Google. (2015a) Where Trends data comes from. Retrieved from:  
[https://support.google.com/trends/answer/4355213?hl=en&ref\\_topic=4365599  
&vid=1-635788917220664688-2248983495](https://support.google.com/trends/answer/4355213?hl=en&ref_topic=4365599&vid=1-635788917220664688-2248983495)
- Google. (2015b) How Trends data is adjusted. Retrieved from  
[https://support.google.com/trends/answer/4365533?hl=en&ref\\_topic=4365599  
&vid=1-635788917220664688-2248983495](https://support.google.com/trends/answer/4365533?hl=en&ref_topic=4365599&vid=1-635788917220664688-2248983495)

M 10/11/16 6:33 PM

Formatted: Space Before: 0 pt

M 10/11/16 6:33 PM

Formatted: Spanish, Check spelling and  
grammar

- Internet Live Stats. (2016, 07 01). *Internet Users by Country (2016)*. Retrieved 09 20, 2016, from Internet Live Stats: <http://www.internetlivestats.com/internet-users-by-country/>
- Investopedia. (2015a). Investopedia. Retrieved from <http://www.investopedia.com/terms/r/return.asp>
- Investopedia. (2015b). Investopedia. Retrieved from <http://www.investopedia.com/terms/v/volume.asp>
- Hanke, J. E., & Reitsch, A. G. (1996). *Pronósticos en los negocios*. 5ta Edición, México:Prentice Hall
- Maletic, J. I., & Marcus, A. (2005). *Data Mining and Knowledge Discovery Handbook*. *Data Mining and Knowledge Discovery Handbook* (pp. 21–36). doi:10.1007/0-387-25465-x\_2
- Market Share Statistics for Internet Technologies. (2015). Netmarketshare.
- MILA. (2015). Mercado Integrado Latinoamericano. Retrieved from <http://www.mercadomila.com/QuienesSomos>
- MinHacienda. (2013). Indicadores ISAAC: Siguiendo la actividad sectorial a partir de Google Trends. *Notas Fiscales*, 1(22). Retrieved from [http://www.minhacienda.gov.co/portal/page/portal/HomeMinhacienda/politicaifiscal/reportesmacroeconomicos/NotasFiscales/22 Siguiendo la actividad sectorial a partir de Google Trends.pdf](http://www.minhacienda.gov.co/portal/page/portal/HomeMinhacienda/politicaifiscal/reportesmacroeconomicos/NotasFiscales/22%20Siguiendo%20la%20actividad%20sectorial%20a%20partir%20de%20Google%20Trends.pdf)
- Moat, H. S., Curme, C., Avakian, A., Kenett, D. Y., Stanley, H. E., & Preis, T. (2013). Quantifying Wikipedia Usage Patterns Before Stock Market Moves. *Scientific Reports*, 3, 1–5. doi:10.1038/srep01801
- Mundial, B. (2015). Banco Mundial. Retrieved from [http://datos.bancomundial.org/indicador/CM.MKT.LCAP.CD?order=wbapi\\_data\\_value\\_2012+wbapi\\_data\\_value+wbapi\\_data\\_value-last&sort=desc](http://datos.bancomundial.org/indicador/CM.MKT.LCAP.CD?order=wbapi_data_value_2012+wbapi_data_value+wbapi_data_value-last&sort=desc)
- NASDAQ. (2015). NASDAQ. Retrieved from <http://www.nasdaq.com/investing/glossary/t/trading-volume>
- Preis, T., Moat, H. S., & Stanley, H. E. (2013). Quantifying trading behavior in financial markets using Google Trends. *Scientific Reports*, 3, 1684. doi:10.1038/srep01684
- Preis, T., Reith, D., & Stanley, H. E. (2010). Complex dynamics of our economic life on different scales: insights from search engine query data. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, 368(1933), 5707–5719. doi:10.1098/rsta.2010.0284
- Rankia. (2012). Rankia Mexico. Retrieved from <http://www.rankia.mx/blog/analisis-ipc/1610753-indices-principales-bmv>



- Simon, H. (1957). A Behavioral Model of Rational Choice. *Models of Man*, 69(1), 99–118.
- Statista. (2015). Statista.
- Statistic Brain Research Institute. (2015). Statistic Brain Research Institute.
- Stats, I. W. (2015). Internet World Stats.
- Telesemana. (2015). TeleSemana. Retrieved from <http://www.telesemana.com/blog/2015/02/16/estadisticas-internet-en-america-latina-solo-chile-superara-el-70-de-penetracion-en-2018/>
- US Media Consulting. (2014). US Media Consulting. Retrieved from <http://usmediaconsulting.com/img/uploads/pdf/El-mercado-de-medios-de-Latinoamrica-2014.pdf>
- Walpole R., Myers R., Myers S., Ye K. Probabilidad y Estadística para ingeniería y ciencias. Octava Edición. Pearson, Prentice Hall. 2007
- World Bank. (2013). *GINI index (World Bank estimate)*. Retrieved 09 20, 2016, from The World Bank : [http://data.worldbank.org/indicator/SI.POV.GINI?year\\_high\\_desc=false](http://data.worldbank.org/indicator/SI.POV.GINI?year_high_desc=false)
- Wu, L., & Brynjolfsson, E. (2009). The Future of Prediction: How Google Searches Foreshadow Housing Prices and Sales. *SSRN Electronic Journal*, 1–24. doi:10.2139/ssrn.2022293
- Zachte, E. (2015). Wikimedia. Retrieved from <http://stats.wikimedia.org/EN/TablesPageViewsMonthlyCombined.htm>
- Harvey, C. R. (2011). *Nasdaq*. Obtenido de Nasdaq.com: <http://www.nasdaq.com/investing/glossary/s/short-position>
- Harvey, C. R. (2011). *Nasdaq*. Obtenido de Nasdaq.com: <http://www.nasdaq.com/investing/glossary/l/long-position>
- BBVA. (10 de 12 de 2015). *BBVA*. Obtenido de BBVA: <https://info.bbva.com/es/noticias/economia/finanzas/invertir-en-bolsa-desde-cero-que-es-la-volatilidad/>



## ANEXO 1

ACTIVOS SELECCIONADOS PARA COLOMBIA					
Ticker	Nombre	Google Trends		Wikipedia	
		Tipo	Término de búsqueda o Tema	Nombre de artículo en inglés	Nombre de artículo en español
GRUPOSUA	Grupo de Inversiones Suramericana SA	Término de búsqueda	Sura (Colombia)	Grupo Sura	Grupo Sura
ECOPETL	Ecopetrol SA	Tema: Company	Ecopetrol	Ecopetrol	Ecopetrol
NUTRESA	Grupo Nutresa SA	Tema: Food processing Company	Grupo Nutresa	Grupo Nutresa	Grupo Nutresa
GRUPOARG	Inversiones Argos SA	Tema: Conglomerate Company	Grupo Argos	Argos S.A.	Cementos Argos
PFAVAL	Grupo Aval Acciones y Valores	Tema: Holding Company	Grupo Aval Acciones y Valores	Grupo Aval Acciones y Valores	Grupo Aval Acciones y Valores
BCOLO	BanColombia SA	Tema: Commercial Banking Company	Bancolombia	Bancolombia	Bancolombia
CORFICOL	Corporación Financiera Colombiana SA	Tema: Company	Corficolombiana	Corficolombiana	-
ISAGEN	Isagen SA ESP	Término de búsqueda	Isagen	Isagen	-
ÉXITO	Almacenes Éxito SA	Tema: Retail Company	Grupo Exito	Grupo Éxito	Almacenes Éxito
BOGOTA	Banco de Bogotá SA	Tema: Commercial Banking Company	Banco de Bogotá	Banco de Bogotá	Banco de Bogotá
PFDVVND	Banco Davivienda SA	Tema: Banking Company	Davivienda	Davivienda	Davivienda
CLH	Cemex Latam Holdings SA	Tema: Manufacturing Company	Cemex	Cemex	Cemex
PFAVH	Avianca Holdings SA	Tema: Airline Company	Avianca	Avianca Holdings	Avianca Holdings
BVC	Bolsa de Valores de Colombia	Tema: Stock exchange	Colombia Stock Exchange	Colombia Stock Exchange	Bolsa de Valores de Colombia
PREC	Pacific Exploration	Tema: Energy	Pacific	Pacific	Pacific

**Relación entre los patrones de búsqueda de internet y los movimientos del MILA**

	and Production Corp	Company	Exploration and Production	Exploration & Production	Exploration & Production
--	---------------------	---------	----------------------------	--------------------------	--------------------------

ACTIVOS SELECCIONADOS PARA MEXICO					
Ticker	Nombre	Google Trends		Wikipedia	
		Tipo	Término de búsqueda o Tema	Nombre de artículo en inglés	Nombre de artículo en español
FEMSAUBD	Fomento Económico Mexicano SAB de CV	Tema: Retail Company	FEMSA	FEMSA	Fomento Económico Mexicano
AMXL	América Móvil SAB de CV	Tema: Telecommunications Company	América Móvil	América Móvil	América Móvil
GFNORTEO	Grupo Financiero Banorte SAB de CV	Tema: Banking Company	Banorte	Banorte	Grupo Financiero Banorte
TLEVICPO	Grupo Televisa SA	Tema: Mass Media Company	Televisa	Televisa	Televisa
WALMEX*	Wal-Mart de México SAB de CV	Tema: Retail Company	Walmart de México y Centroamérica	Walmart de México y Centroamérica	Walmart de México y Centroamérica
GMEXICOB	Grupo México SAB de CV	Mining Company	Grupo México	Grupo México	Grupo México
CEMEXCPO	Cemex SAB de CV	Tema: Manufacturing Company	Cemex	Cemex	Cemex
ALFAA	Alfa SAB de CV	Tema: Multinational Conglomerate Company	Grupo Industrial Alfa S.A.B de C.V	Alfa (Mexico)	Alfa (Empresa de México)
GAPB	Grupo Aeroport-B	Tema: Airport Operator	Grupo Aeroportuario del Pacífico	Grupo Aeroportuario del Pacífico	Grupo Aeroportuario del Pacífico
KOFL	Coca-Cola Femsa SAB de CV	Tema: Company	Coca Cola Femsa S.A.B. de C.V.	Coca Cola Femsa	Coca Cola Femsa
KIMBERA	Kimberly-Clark de México SAB de CV	Tema: Corporation	Kimberly-Clark	Kimberly-Clark de México	Kimberly-Clark de México

Relación entre los patrones de búsqueda de internet y los movimientos del MILA

BIMBOA	Grupo Bimbo SAB de CV	Tema: Company	Grupo Bimbo	Grupo Bimbo	Grupo Bimbo
GRUMAB	Gruma SAB de CV	Tema: Manufacturing Company	Gruma	Gruma	Gruma
GFINBURO	Grupo Financiero Inbursa SA	Tema: Company	Inbursa	Inbursa	Grupo Financiero Inbursa
LIVEPOLC	El Puerto de Liverpool SAB de CV	Tema: Retail Company	Liverpool	Liverpool (store)	El Puerto de Liverpool
MEXCHEM*	Mexichem SAB de CV	Tema: Company	Mexichem	Mexichem	-
GENTERA*	Gentera SAB de CV	Tema: Bank	Compartamos banco	Compartamos banco	-
ELEKTRA*	Grupo Elektra SAB de CV	Tema: Company	Grupo Salinas	Grupo Elektra	Grupo Elektra
AC*	Arca Continental SAB de CV	Tema: Manufacturing Company	Arca Continental	Arca Continental	-
ALSEA*	Alsea SAB de CV	Tema: Company	Alsea Restaurant	Alsea (company)	Alsea (empresa)
LALAB	Grupo Lala SAB de CV	Tema: Company	Grupo Lala	Grupo Lala	Grupo Lala
GCARSOA1	Grupo Carso SAB de CV	Tema: Conglomerate Company	Grupo Carso	Grupo Carso	Grupo Carso
PE&OLES*	Industrias Peñoles SAB de CV	Tema: Company	Peñoles Mining	Peñoles	Industrias Peñoles
IENOVA*	Infraestructura Energetica Nova SAB de C	Término de búsqueda	Ienova	-	IENOVA
OHLMEX*	OHL Mexico SAB de CV	Tema: Construction Company	Obrascón Huarte Lain	Obrascón Huarte Lain	Obrascón Huarte Lain
LABB	Genomma Lab Internacional SAB de CV	Tema: International Business Operation	Genomma Lab	-	Genomma Lab Internacional
GFREGIO	Banregio Grupo Financiero SAB de CV	Tema: Banking Company	BanRegio	BanRegio	BanRegio

**Relación entre los patrones de búsqueda de internet y los movimientos del MILA**

ICA*	Empresas ICA SAB de CV	Tema: Company	Empresas ICA	-	Empresas ICA
SANMEXB	Grupo Financiero Santander SAB de CV	Tema: Bank	Santander Bank	Santander Group	Grupo Financiero Santander

ACTIVOS SELECCIONADOS PARA CHILE					
Ticker	Nombre	Google Trends		Wikipedia	
		Tipo	Término de búsqueda o Tema	Nombre de artículo en inglés	Nombre de artículo en español
ENERSIS	Enersis SA	Tema: Corporation	Enersis S.A	Enersis	Enersis
COPEC	Empresas COPEC SA	Término de Búsqueda	Copec	Empresas Copec	Empresas Copec
ENDESA	Endesa (Chile)	Término de Búsqueda	Endesa	Endesa (Chile)	Endesa (Chile)
FALAB	SACI Falabella	Tema: Department Store Company	Falabella	Falabella	S.A.C.I. Falabella / Falabella (retail store)
BSAN	Banco Santander Chile	Tema: Commercial Banking Company	Banco Santander-Chile	Banco Santander (Chile)	Banco Santander-Chile
CMPC	Empresas CMPC SA	Tema: Company	CMPC	Empresas CMPC	CMPC (company)
CENCOSUD	Cencosud SA	Tema: Retail Company	Cencosud	Cencosud	Cencosud
LAN	Lan Airlines SA	Tema: Airline Company	Lan Airlines	LAN Airlines	LAN Airlines
COLBUN	Colbún SA	Tema: Utility Company	Colbún SA	Colbún S.A.	Colbún S.A.
CHILE	Banco de Chile	Tema: Bank	Banco de Chile	Banco de Chile	Banco de Chile
BCI	Banco de Crédito e Inversiones	Tema: Private Banking Company	Banco de Crédito e Inversiones	Banco de Crédito e Inversiones	Banco de Crédito e Inversiones
SQM/B	Sociedad Química y Minera de Chile SA	Tema: Company	Sociedad Química y Minera	Sociedad Química y Minera de Chile	Sociedad Química y Minera
CORPBANC	Corpbanca	Tema: Bank	Corpbanca	CorpBanca	CorpBanca

**Relación entre los patrones de búsqueda de internet y los movimientos del MILA**

AGUAS/A	Águas Andinas SA	Tema: Business Operation	Águas Andinas	Aguas Andinas	-
SMCHILEB	Sociedad Matriz Banco de Chile	Tema: Bank	Banco de Chile	Banco de Chile	SM-Chile
AESGENER	AES Gener SA	Tema: Company	AES Gener	AES Gener	AES Gener
ANTAR	Antarchile SA	Tema: Conglomerate Company	AntarChile	AntarChile	AntarChile
ENTEL	ENTEL Chile SA	Tema: Telecommunications company	Entel	Entel Chile	Entel (Chile)
PARAUCO	Parque Arauco SA	Tema: Company	Parque Arauco S.A.	Parque Arauco	Parque Arauco S.A.
SONDA	Sonda SA	Tema: Information Technology Company	Sonda S.A.	-	Sonda S.A.
ANDINAB	Embotelladora Andina SA	Tema: Bottled and Canned Soft Drinks and Carbonated Waters Business	Embotelladora Andina	Embotelladora Andina	-
CONCHA	Viña Concha y Toro SA	Tema: Wineries company	Viña Concha y Toro S.A.	Concha y Toro	Concha y Toro
IAM	Inversiones Aguas Metropolitanas SA	Tema: Business Operation	Águas Andinas	-	Inversiones Aguas Metropolitanas
GASCO	Gasco SA	Tema: Limited Company	Abu Dhabi Gas Industries	Gasco	-
RIPLEY	Ripley Corporación SA	Tema: Department store company	Ripley S.A.	Ripley (tienda)	Ripley S.A.
SK	Sigdo Koppers SA	Tema: Conglomerate company	Sigdo Koppers	Sigdo Koppers	Sigdo Koppers
VAPORES	Compañía Sud Americana de Vapores SA	Tema: Company	CSAV	Compañía Sudamericana de Vapores	CSAV
SALFACOR	Salfacorp SA	Tema: Construction company	Salfacorp	SalfaCorp	SalfaCorp

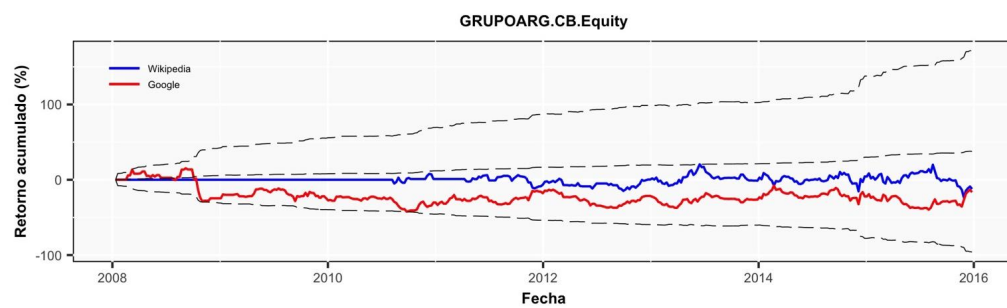
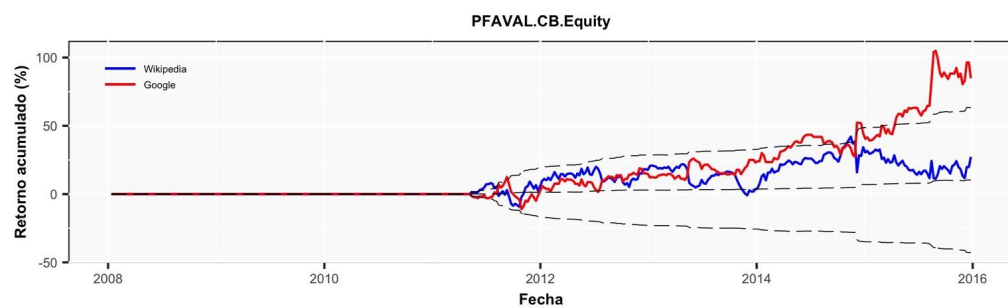
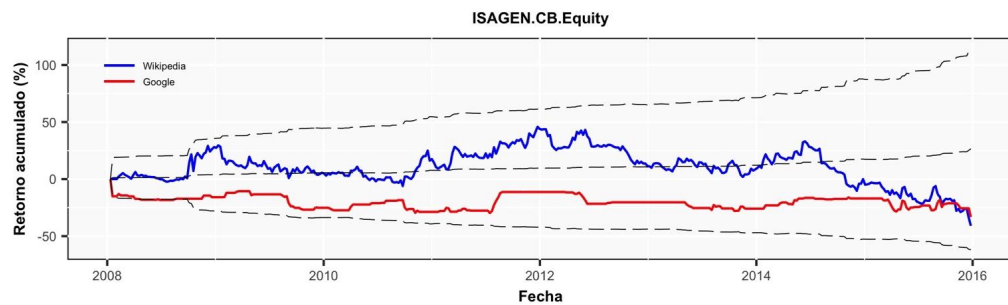
ACTIVOS SELECCIONADOS PARA PERÚ					
Ticker	Nombre	Google Trends		Wikipedia	
		Tipo	Término de	Nombre de	Nombre de

**Relación entre los patrones de búsqueda de internet y los movimientos del MILA**

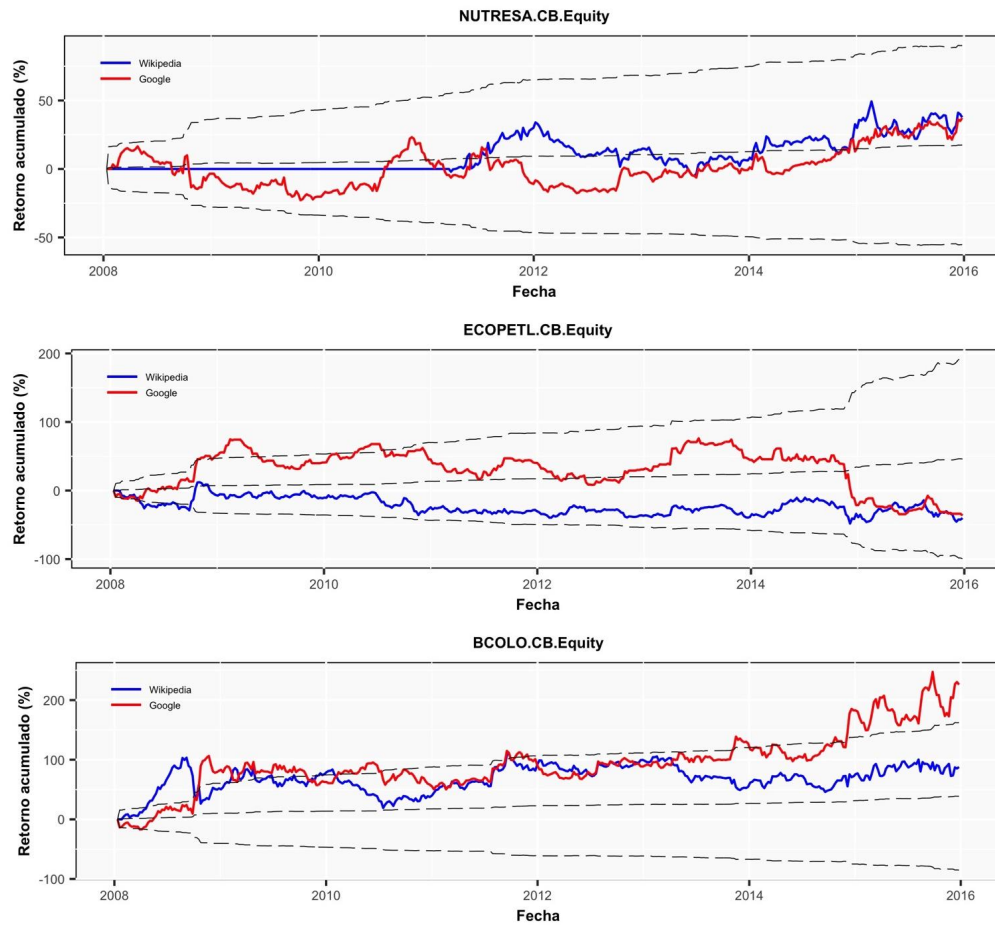
			<b>búsqueda o Tema</b>	<b>artículo en inglés</b>	<b>artículo en español</b>
CORAREC1	Corporación Aceros Arequipa SA	Término de Búsqueda	Aceros Arequipa	Corporación Aceros Arequipa	Corporación Aceros Arequipa
MILPOC1	Compañía Minera Milpo SAA	Tema: Mining Company	Compañía Minera Milpo	Compañía Minera Milpo	-
SCCO	Southern Copper Corp	Tema: Mining Company	Southern Copper Corporation	Southern Copper Corporation	-
ALICORC1	Alicorp SA	Tema: Company	Alicorp	Alicorp	Alicorp
BACKUSI1	Unión Cerv Bac-I	Tema: Brewery	Backus and Johnston Brewery	Backus and Johnston Brewery	Unión de Cervecerías Peruanas Backus y Johnston
IFS	InterCorp Financial Services Corp	Término de Búsqueda	InterCorp	Intergroup Financial Services	Intergroup Financial Services Corp
CVERDEC1	Sociedad Minera Cerro Verde SAA	Mining Company	Cerro Verde	Cerro Verde	Sociedad Minera Cerro Verde
CPACASC1	Cementos Pacasmayo SAA	Cement Company	Cementos Pacasmayo	Cementos Pacasmayo	Pacasmayo
MINSURI1	Minsur SA	Mining Company	Minsur	Minsur	-
LUSURC1	Luz del Sur SAA	Tema: Company	Luz Del Sur	Luz del Sur	Luz del Sur
BVN	Compañía de Minas Buenaventura SA	Tema: Mining company	Buenaventura	Buenaventura (mining Company)	Compañía de Minas Buenaventura



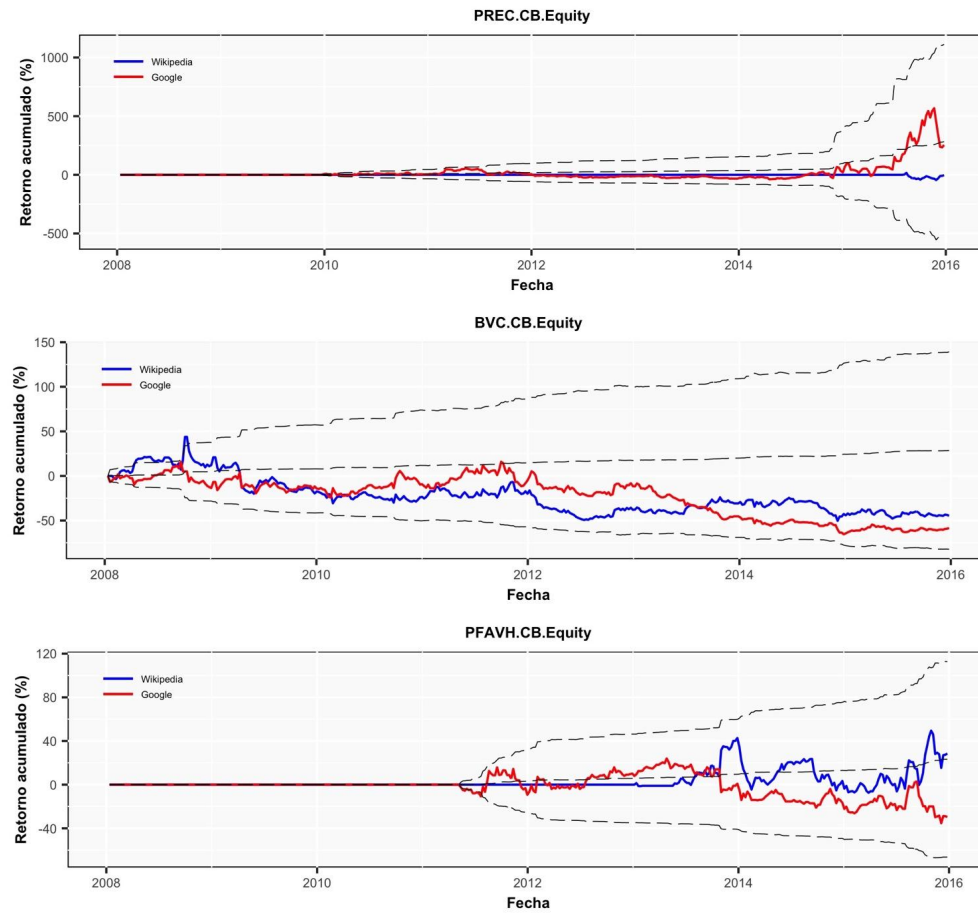
# Relación entre los patrones de búsqueda de internet y los movimientos del MILA



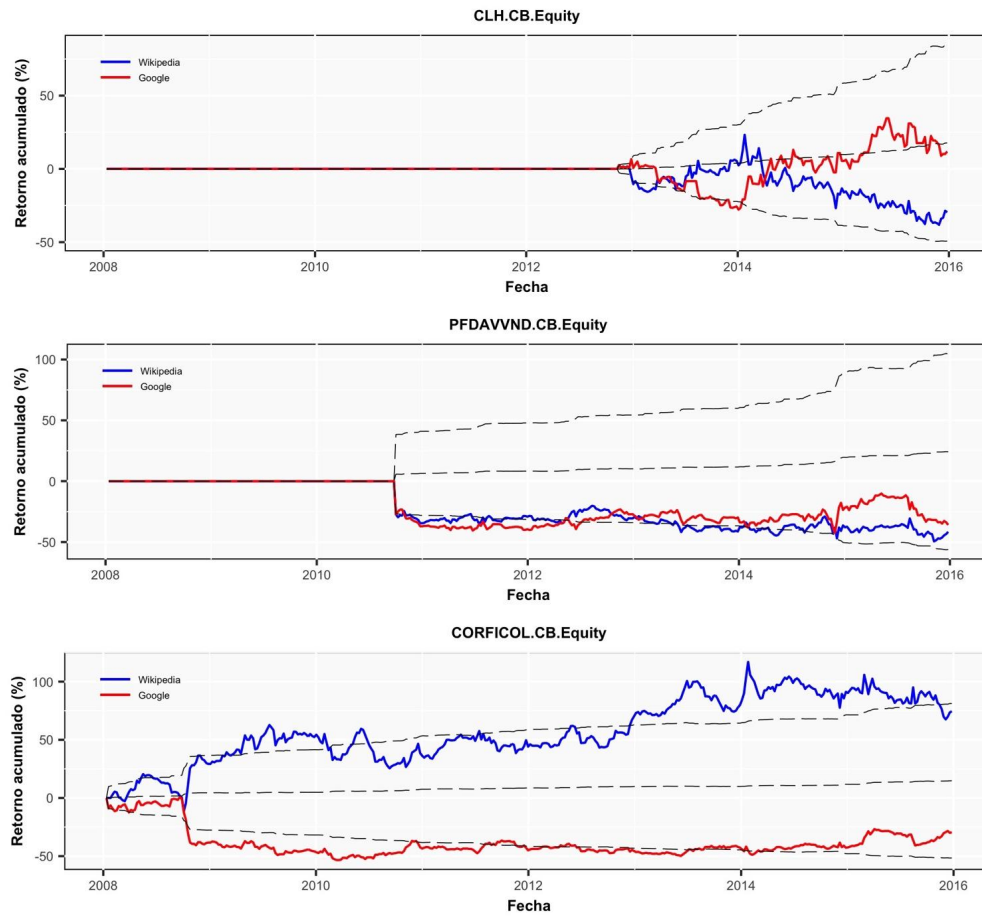
# Relación entre los patrones de búsqueda de internet y los movimientos del MILA



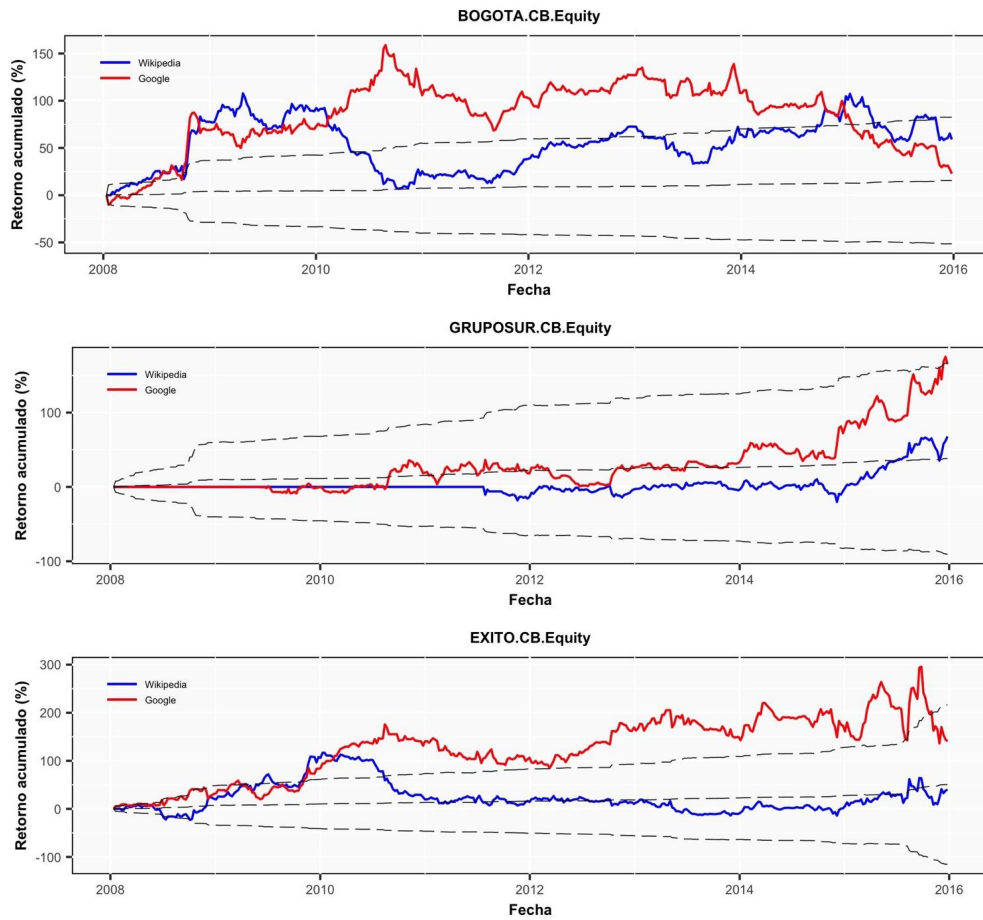
## Relación entre los patrones de búsqueda de internet y los movimientos del MILA



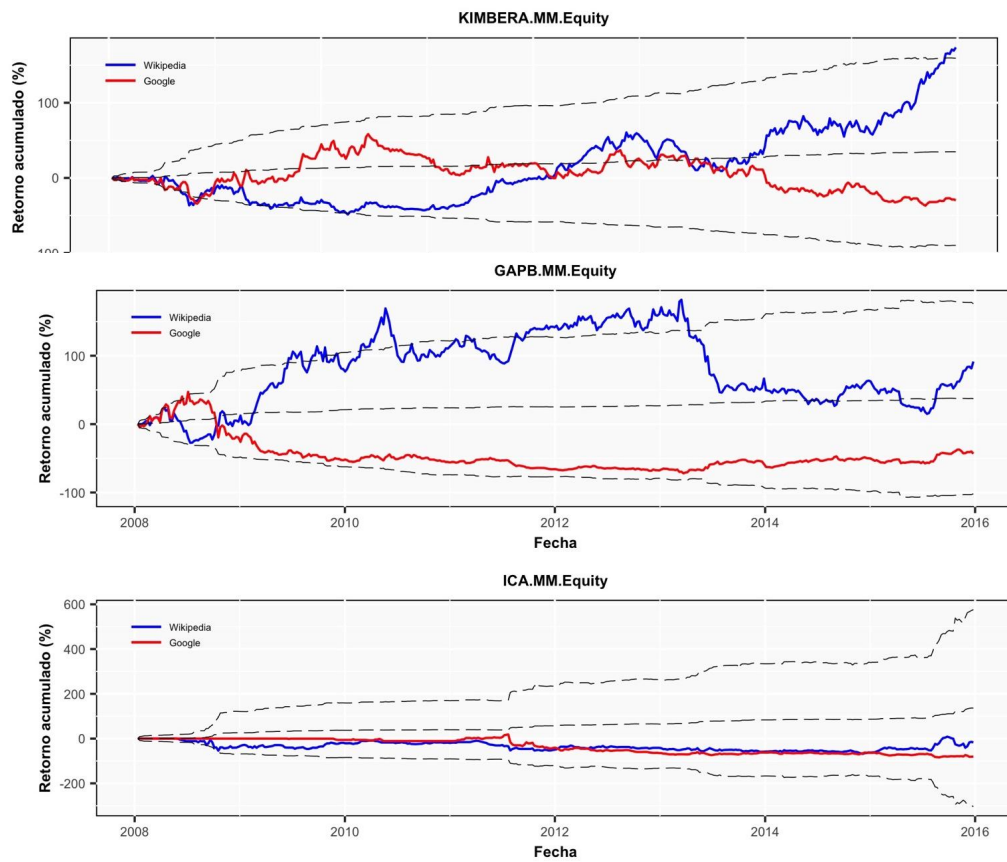
# Relación entre los patrones de búsqueda de internet y los movimientos del MILA



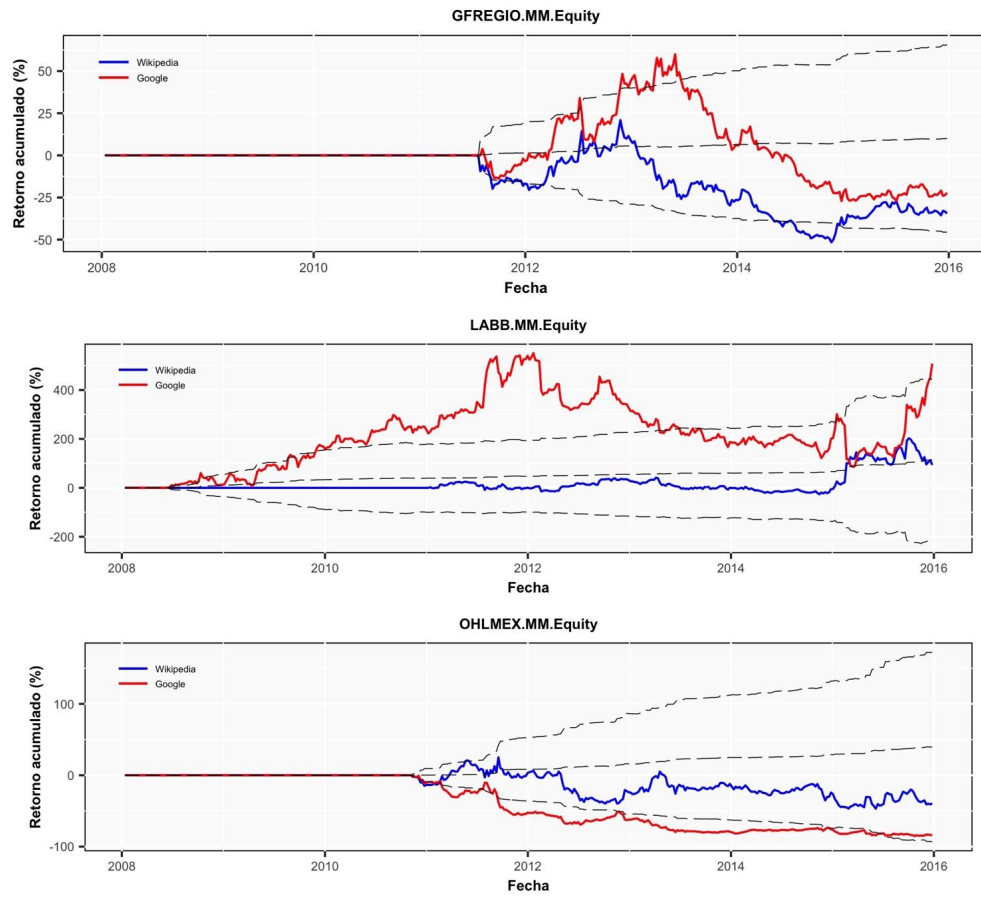
# Relación entre los patrones de búsqueda de internet y los movimientos del MILA



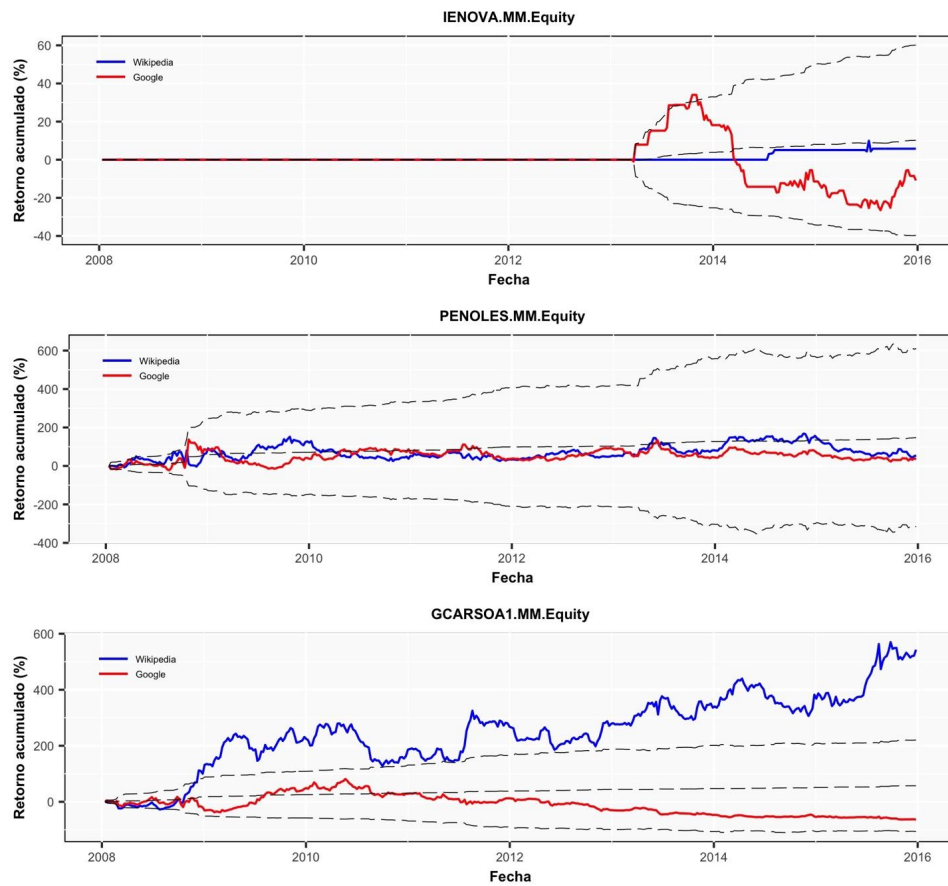
# Relación entre los patrones de búsqueda de internet y los movimientos del MILA



## Relación entre los patrones de búsqueda de internet y los movimientos del MILA

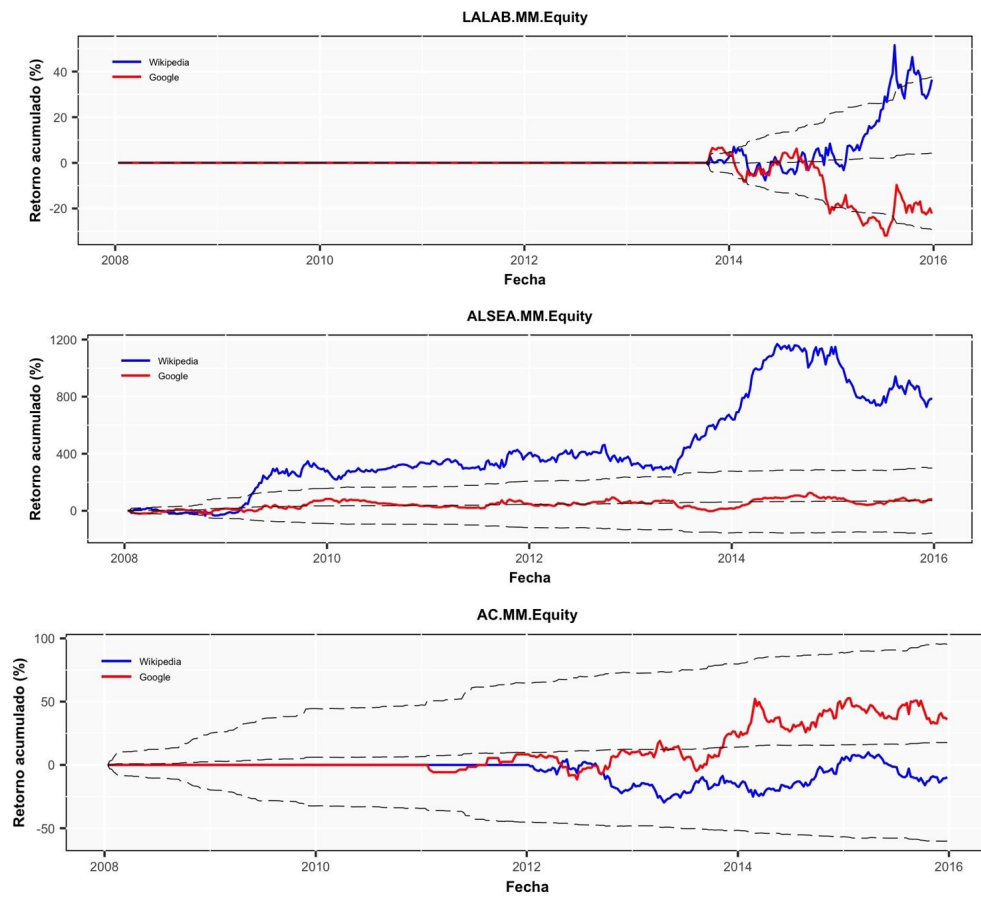


## Relación entre los patrones de búsqueda de internet y los movimientos del MILA

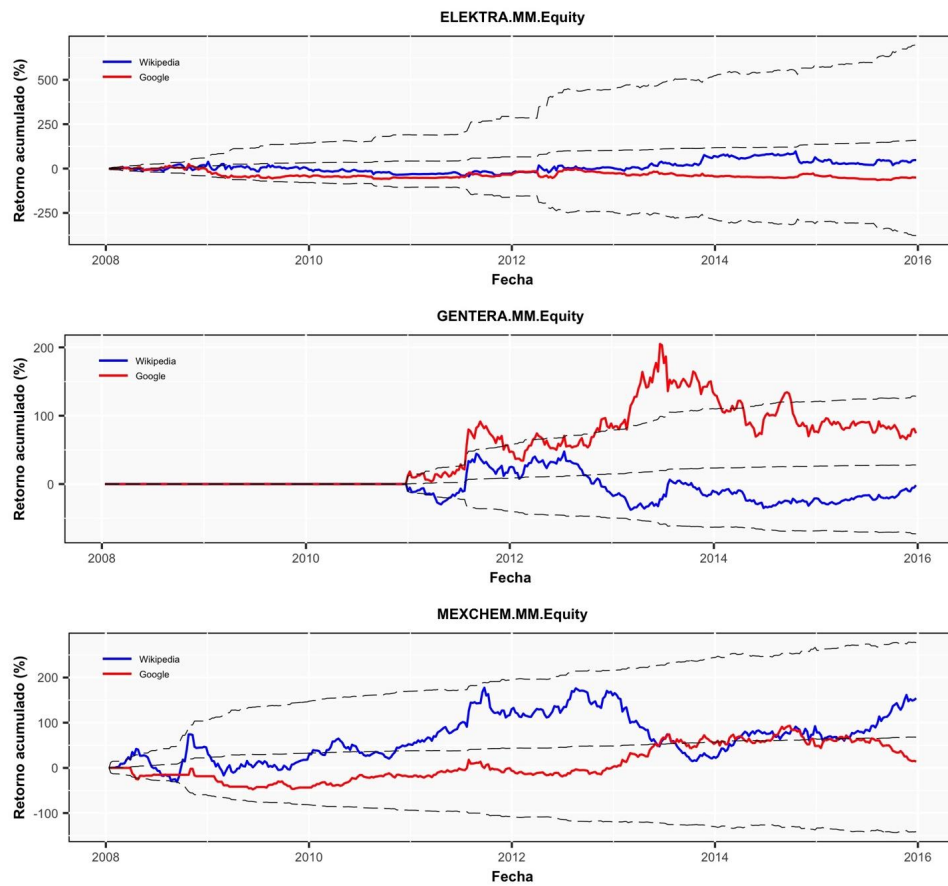




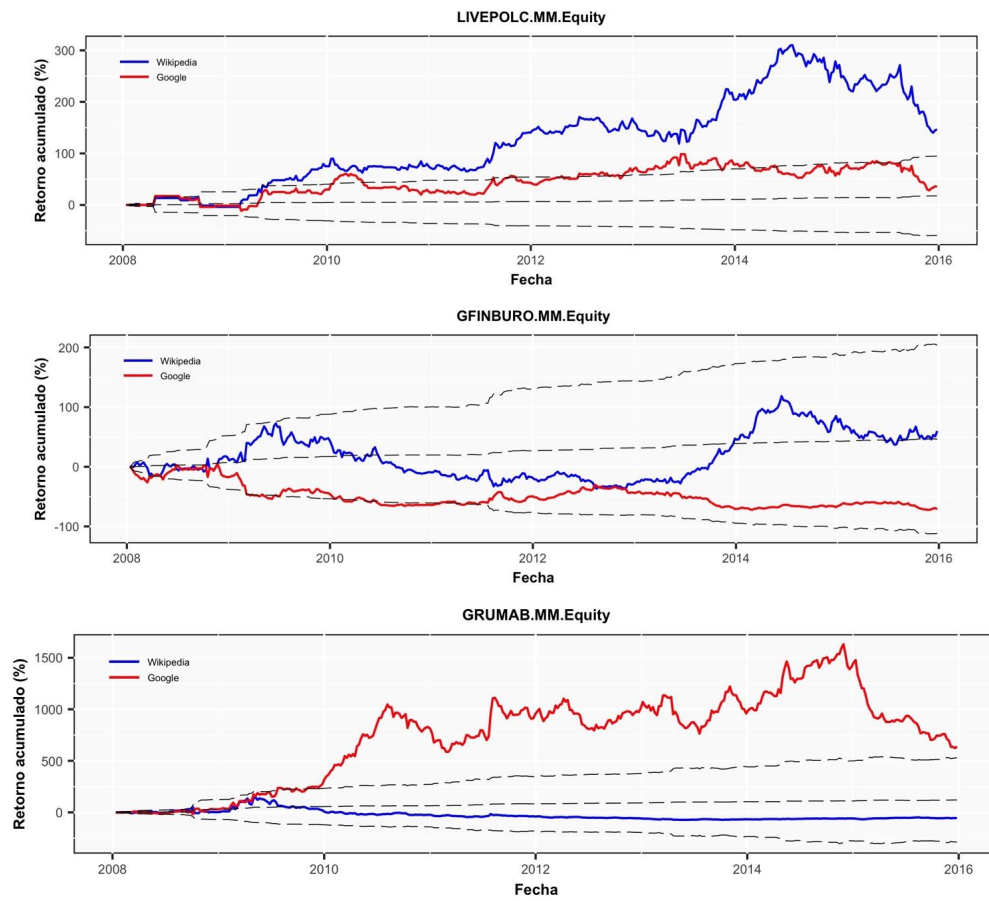
# Relación entre los patrones de búsqueda de internet y los movimientos del MILA



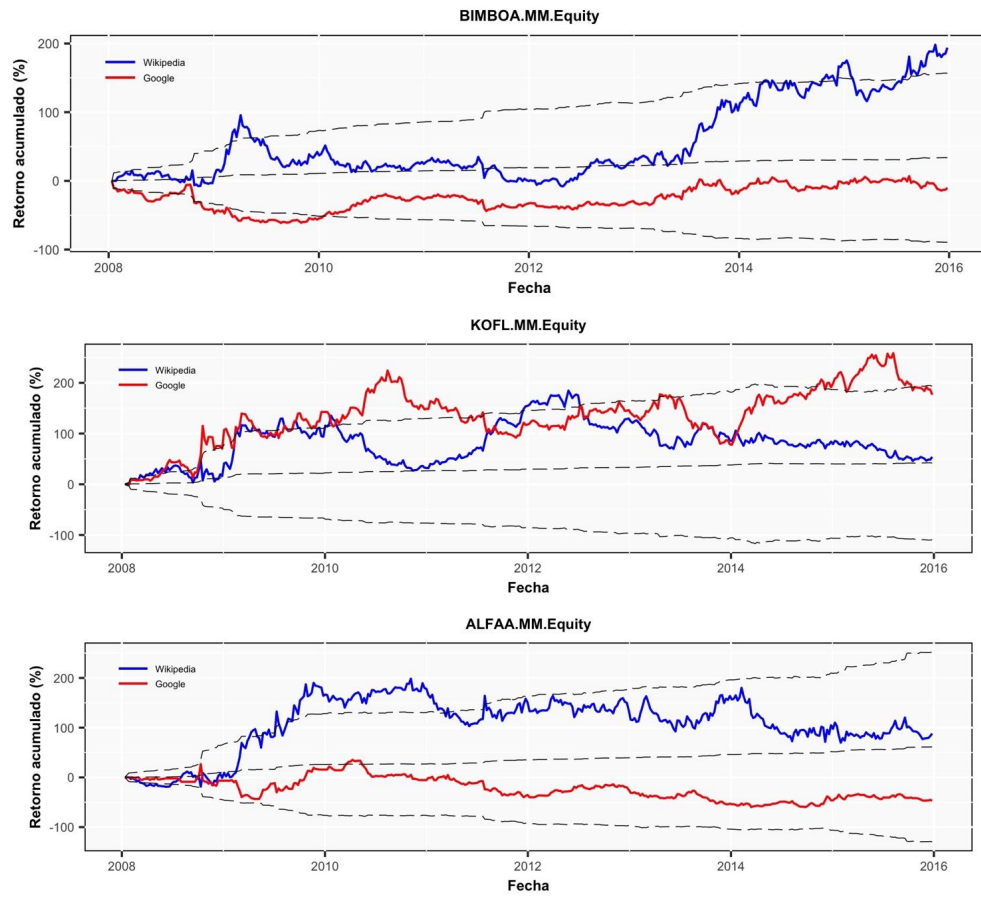
## Relación entre los patrones de búsqueda de internet y los movimientos del MILA



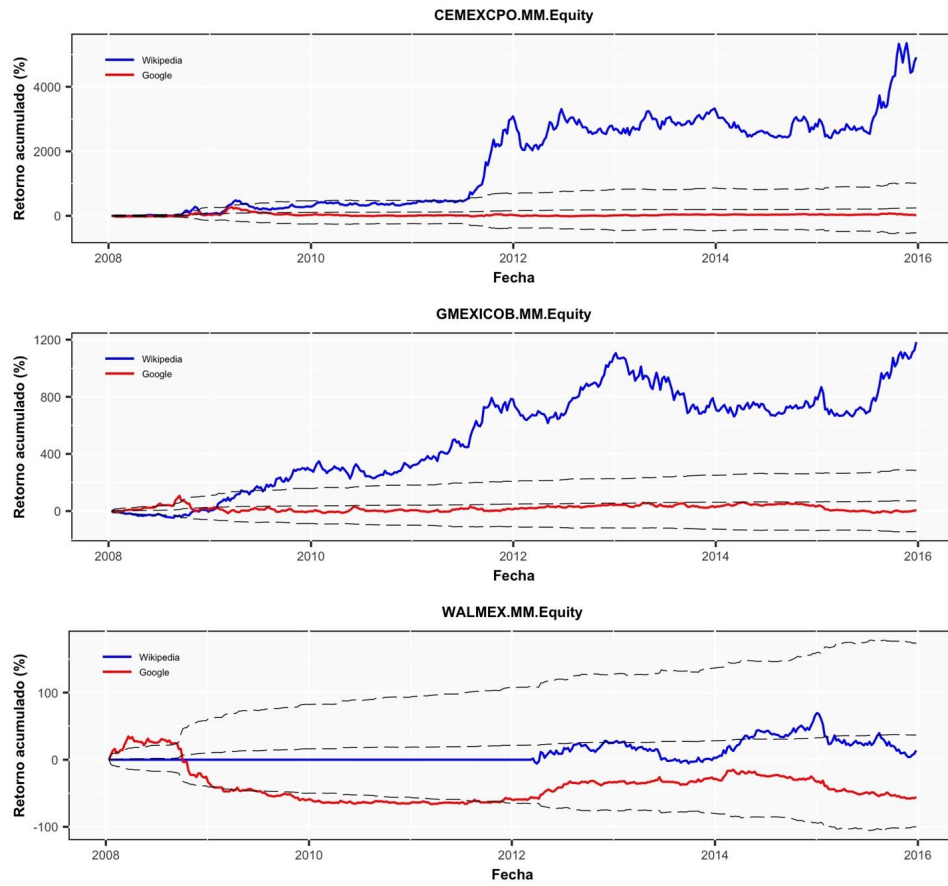
# Relación entre los patrones de búsqueda de internet y los movimientos del MILA



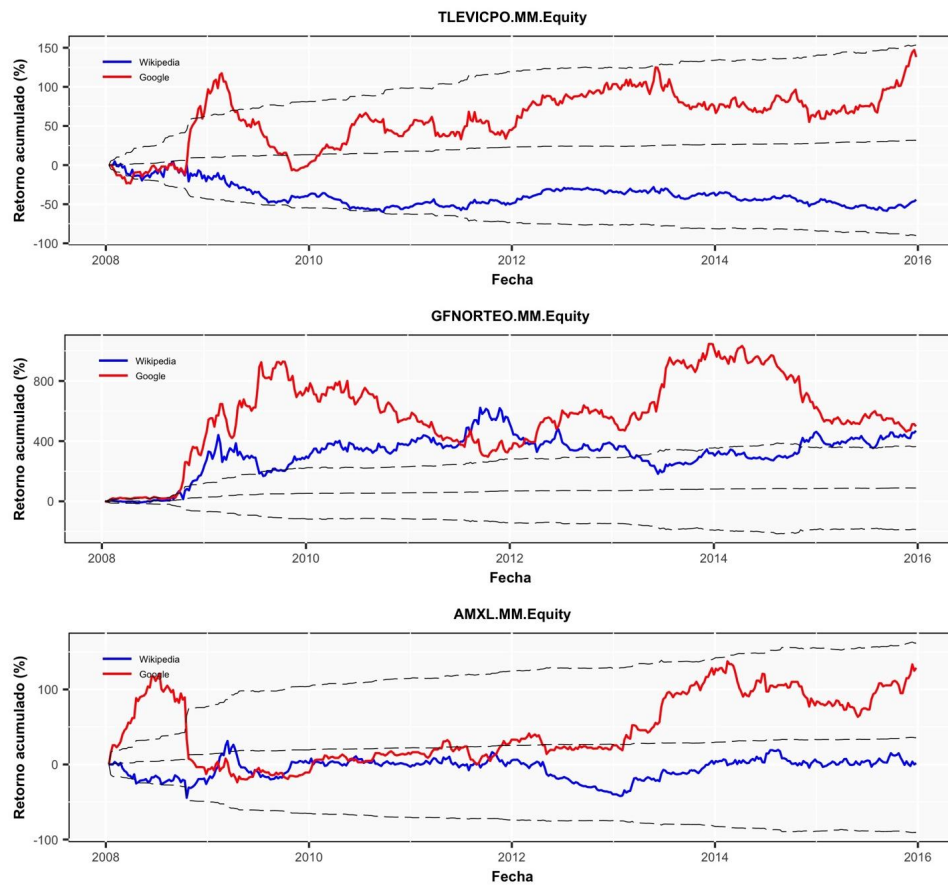
# Relación entre los patrones de búsqueda de internet y los movimientos del MILA



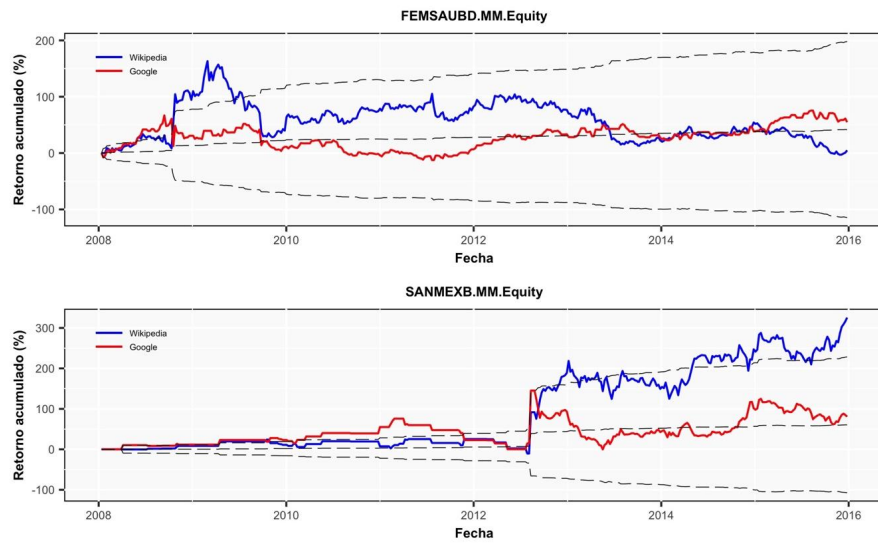
# Relación entre los patrones de búsqueda de internet y los movimientos del MILA



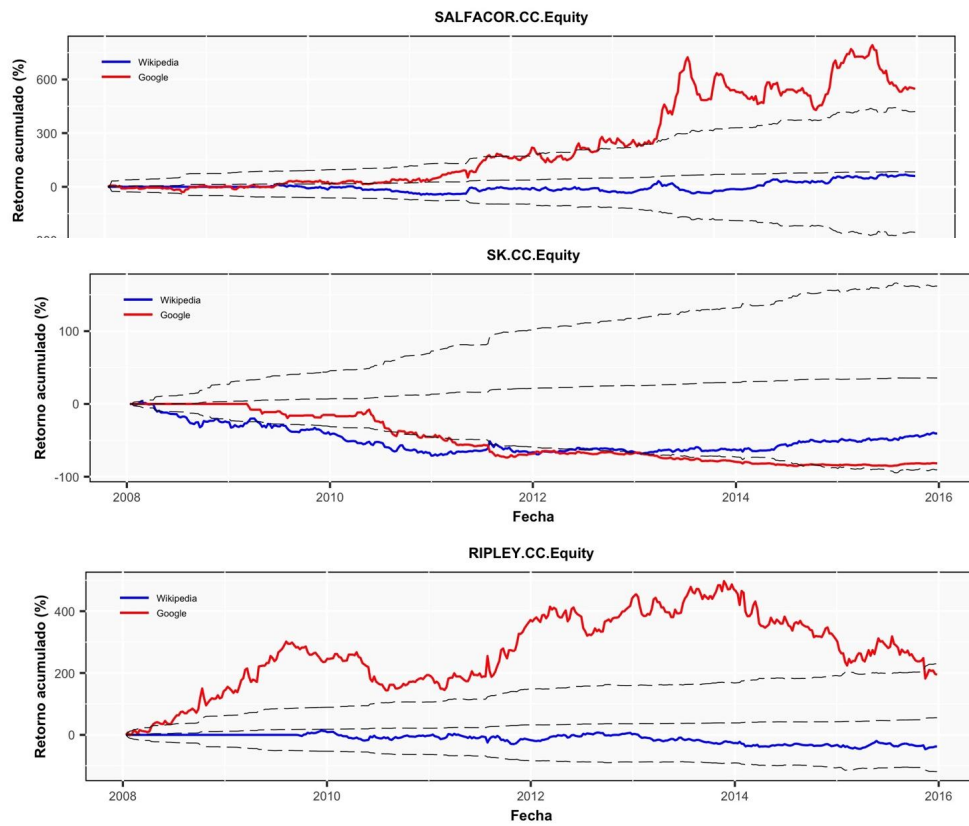
## Relación entre los patrones de búsqueda de internet y los movimientos del MILA



## Relación entre los patrones de búsqueda de internet y los movimientos del MILA

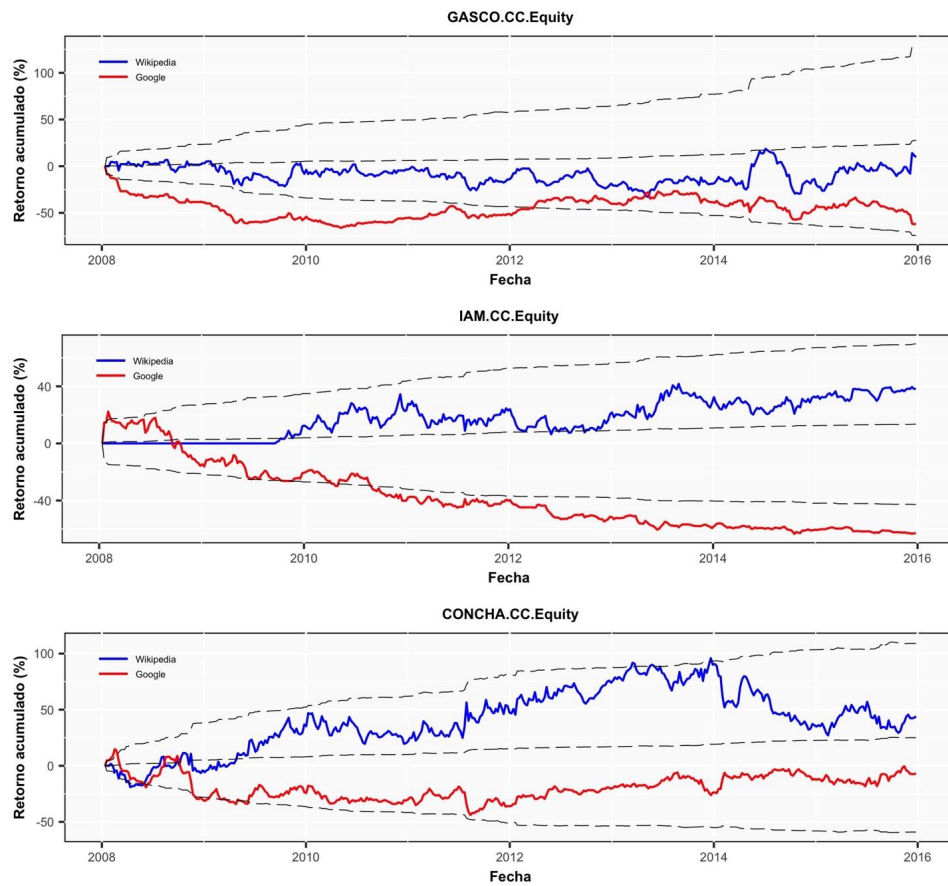


# Relación entre los patrones de búsqueda de internet y los movimientos del MILA

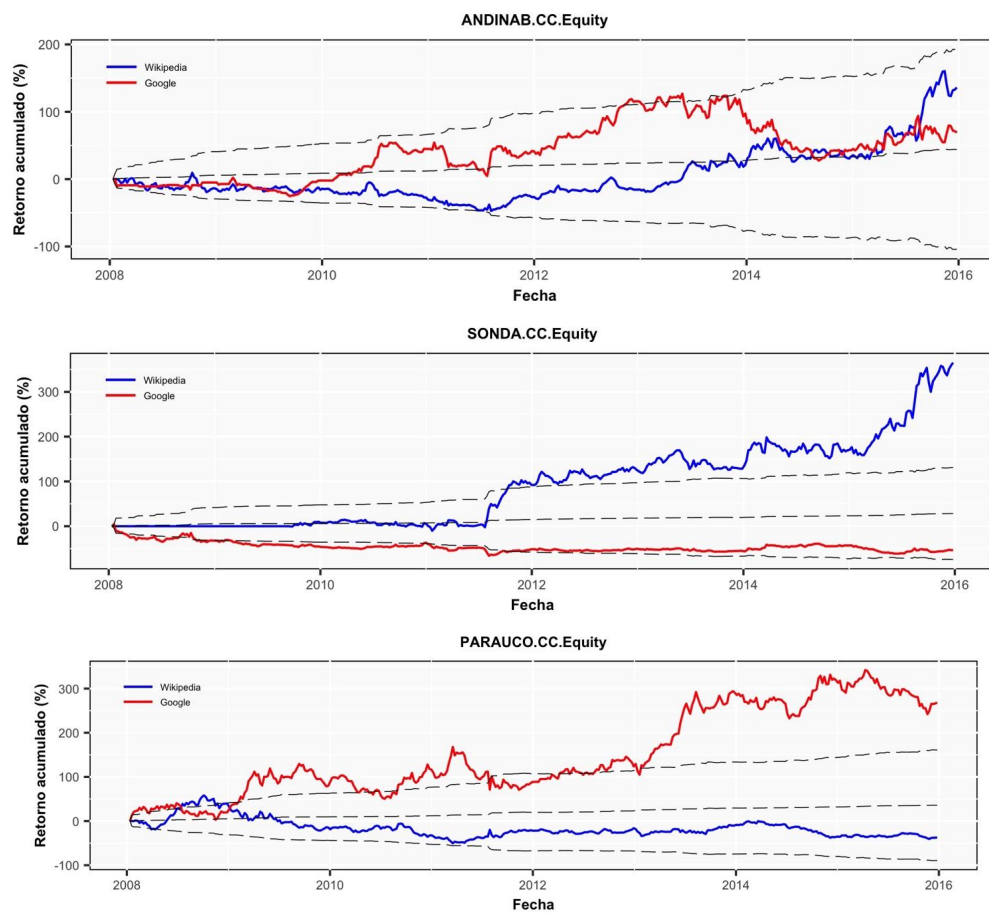




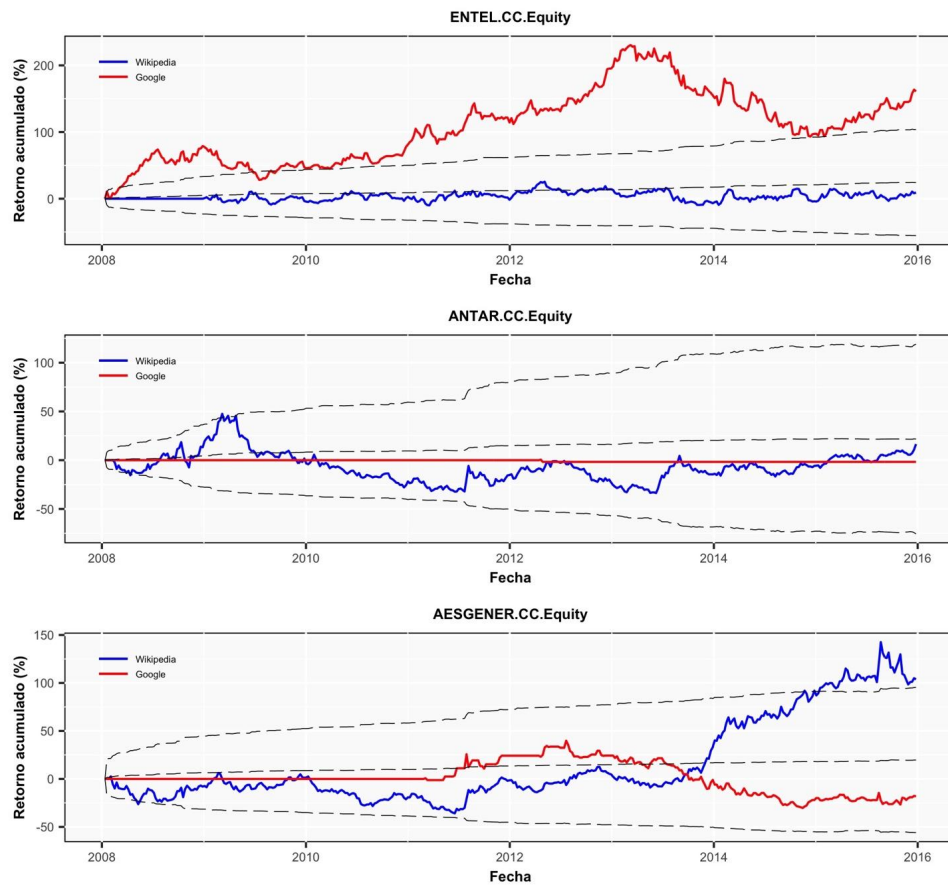
## Relación entre los patrones de búsqueda de internet y los movimientos del MILA



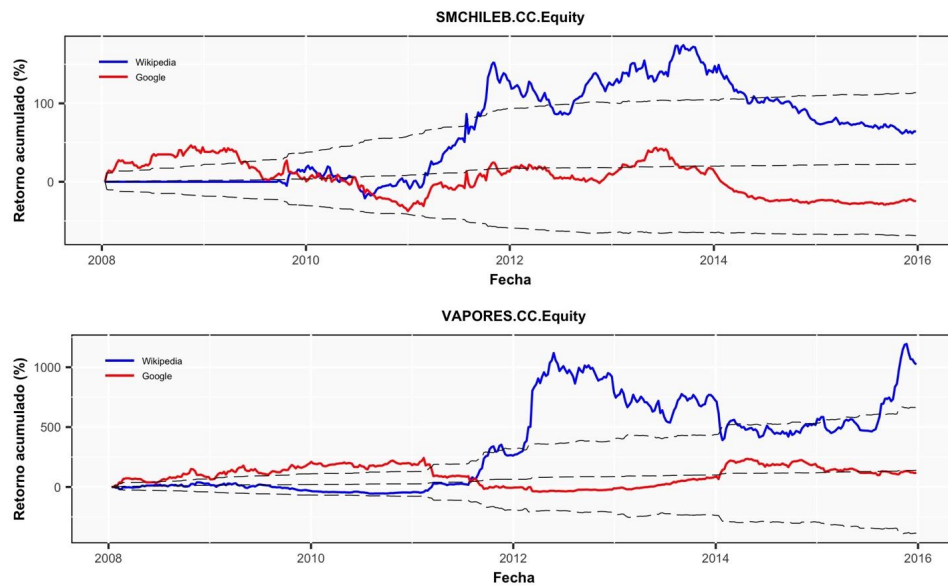
# Relación entre los patrones de búsqueda de internet y los movimientos del MILA



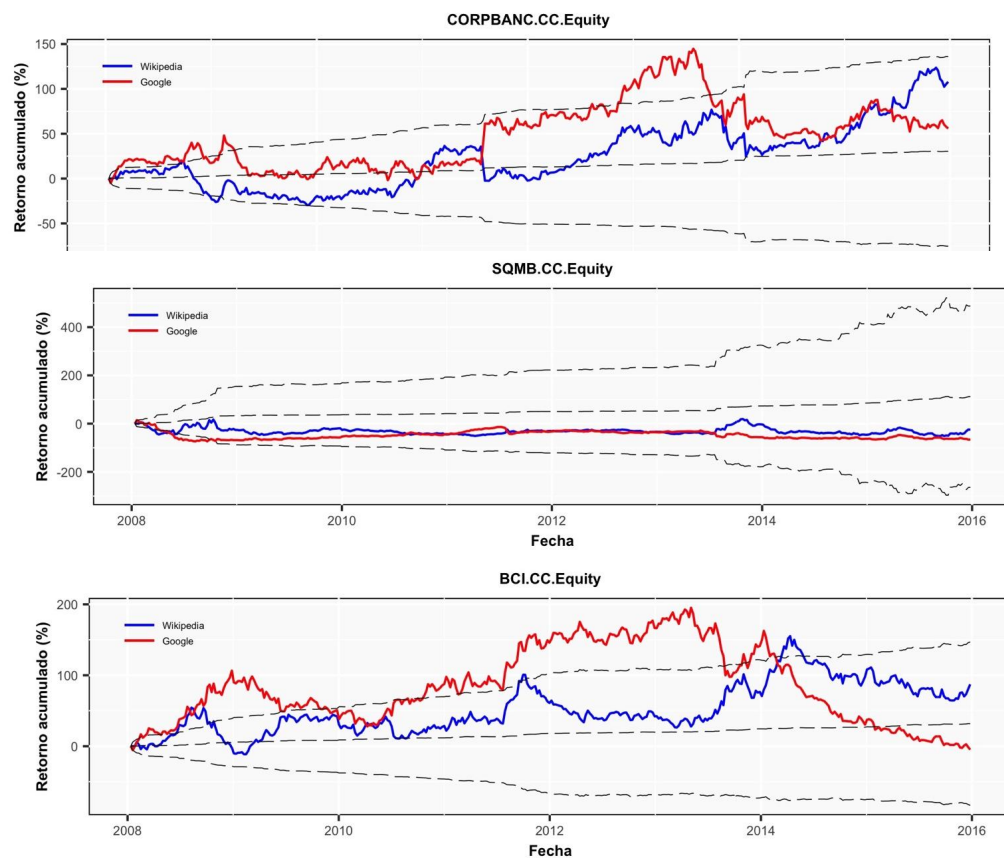
## Relación entre los patrones de búsqueda de internet y los movimientos del MILA



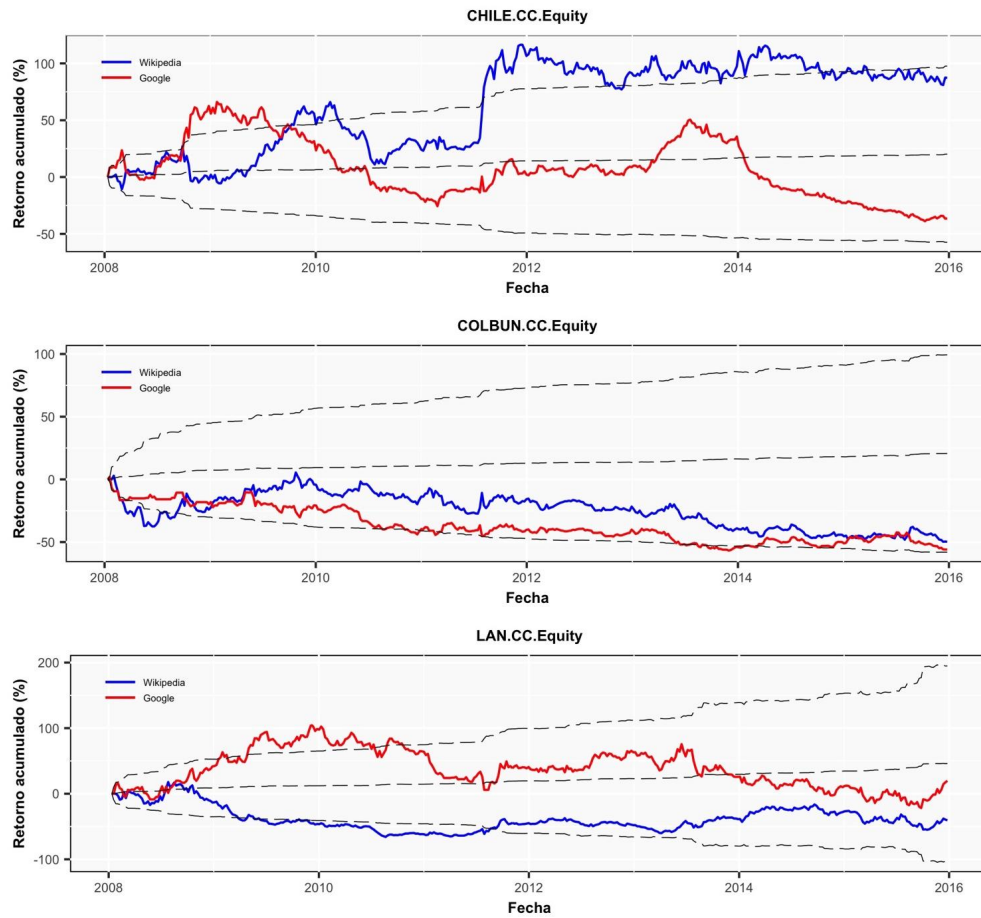
# Relación entre los patrones de búsqueda de internet y los movimientos del MILA



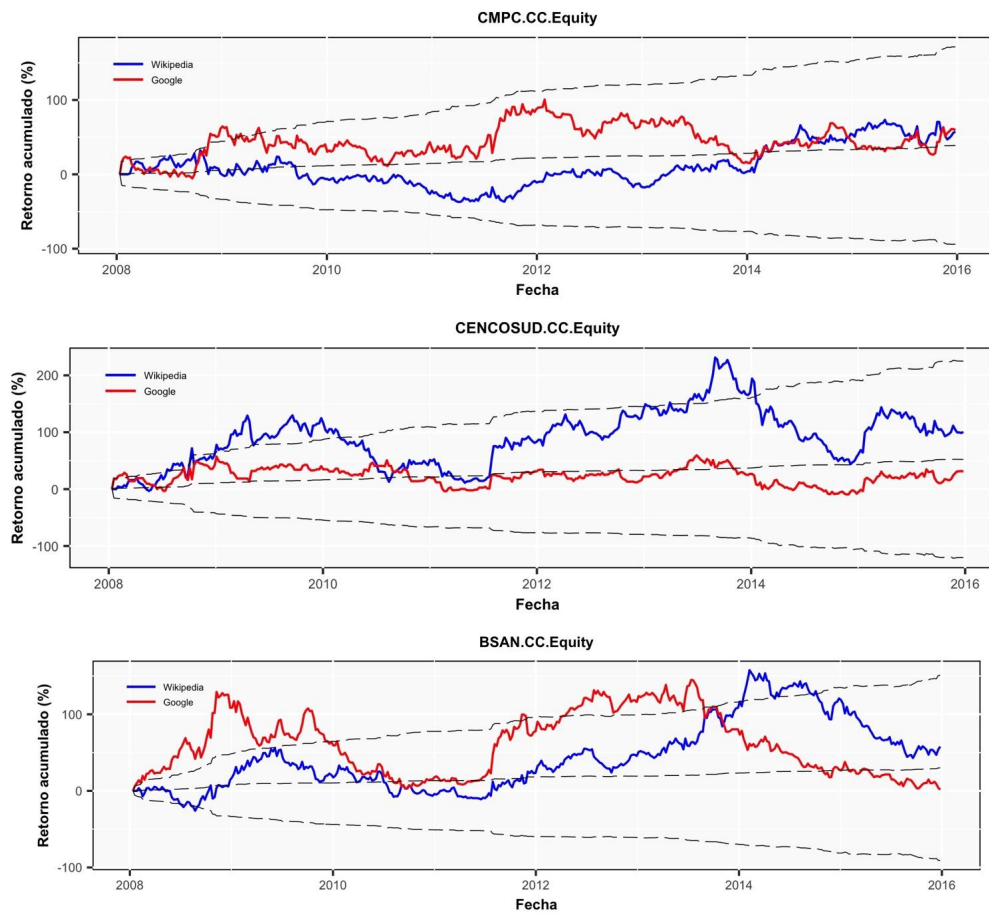
# Relación entre los patrones de búsqueda de internet y los movimientos del MILA



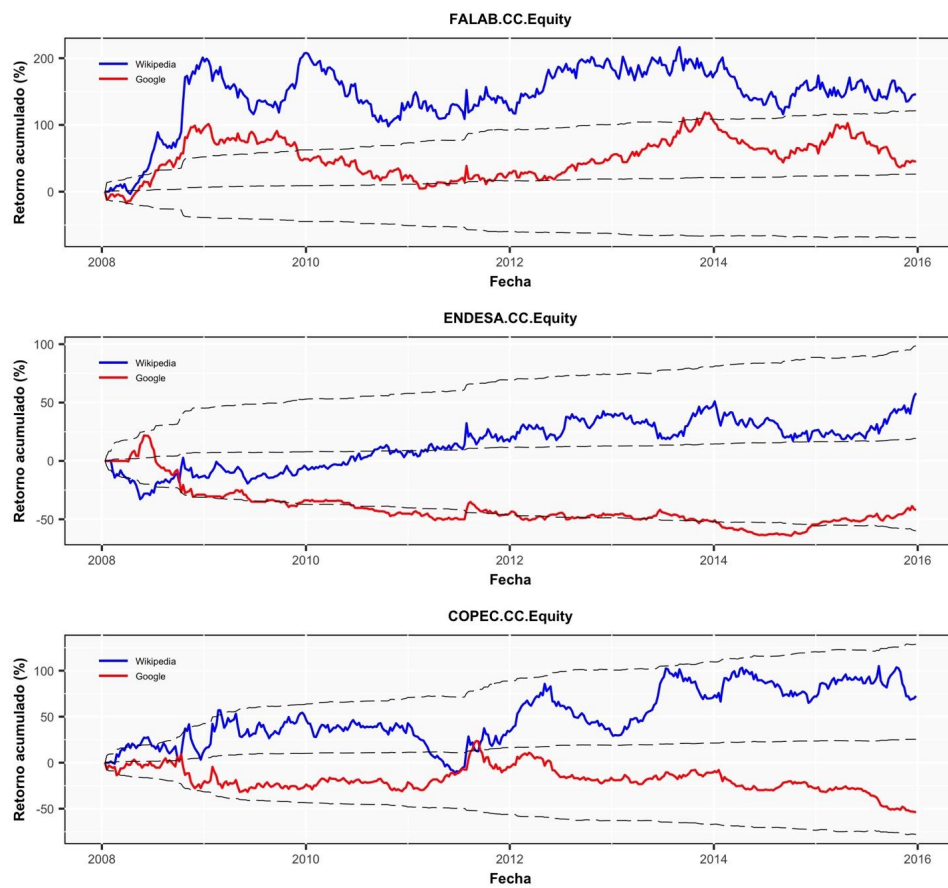
# Relación entre los patrones de búsqueda de internet y los movimientos del MILA



# Relación entre los patrones de búsqueda de internet y los movimientos del MILA

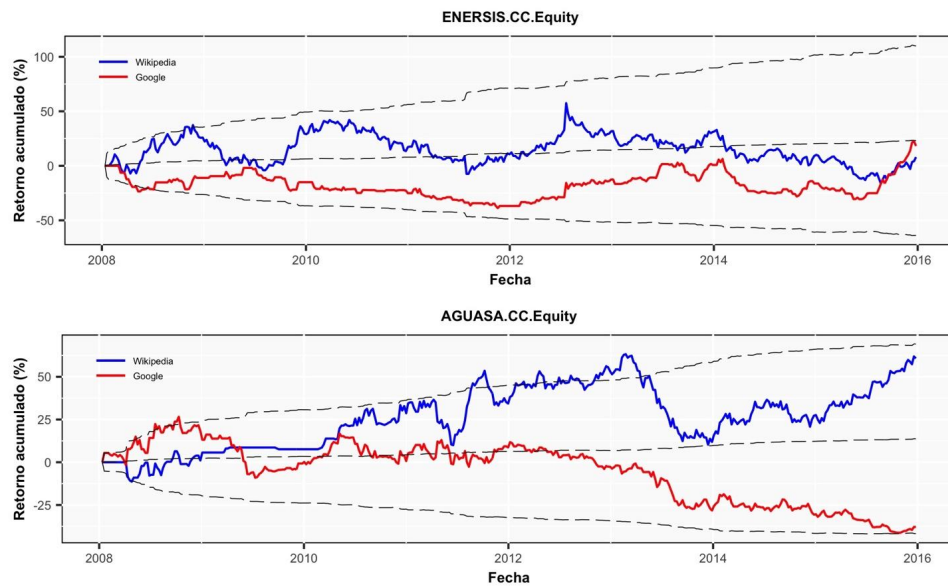


# Relación entre los patrones de búsqueda de internet y los movimientos del MILA

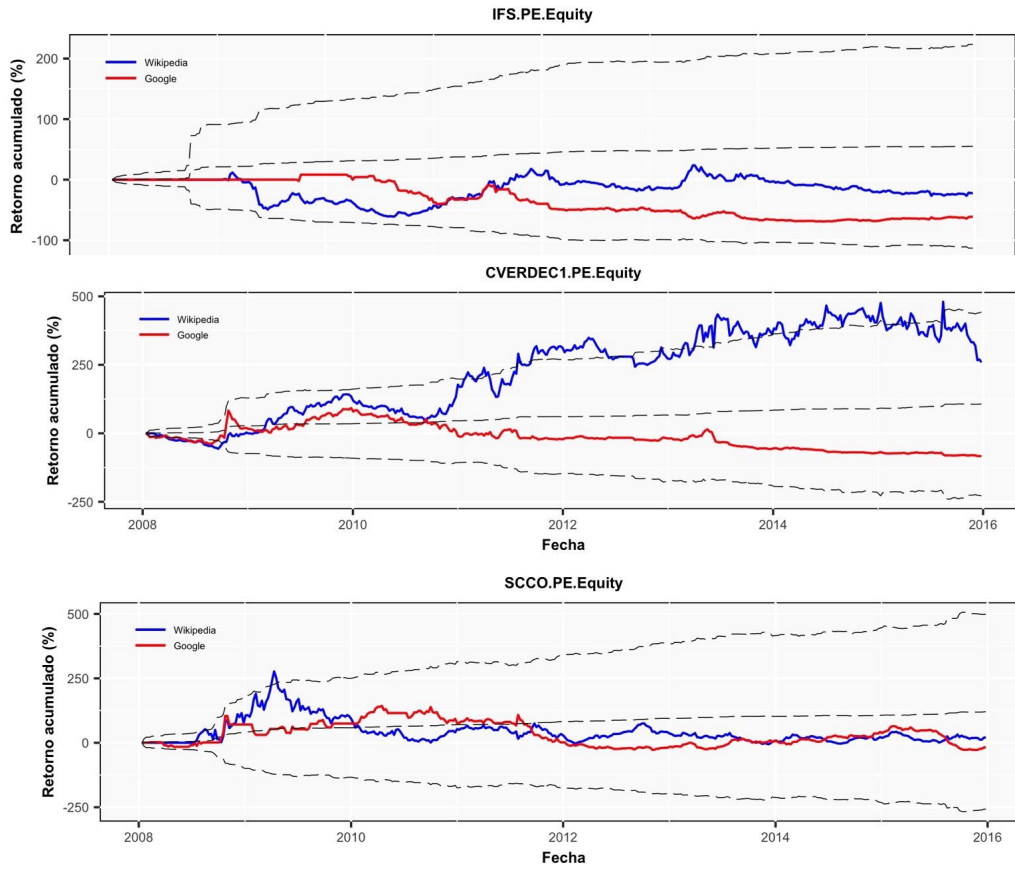




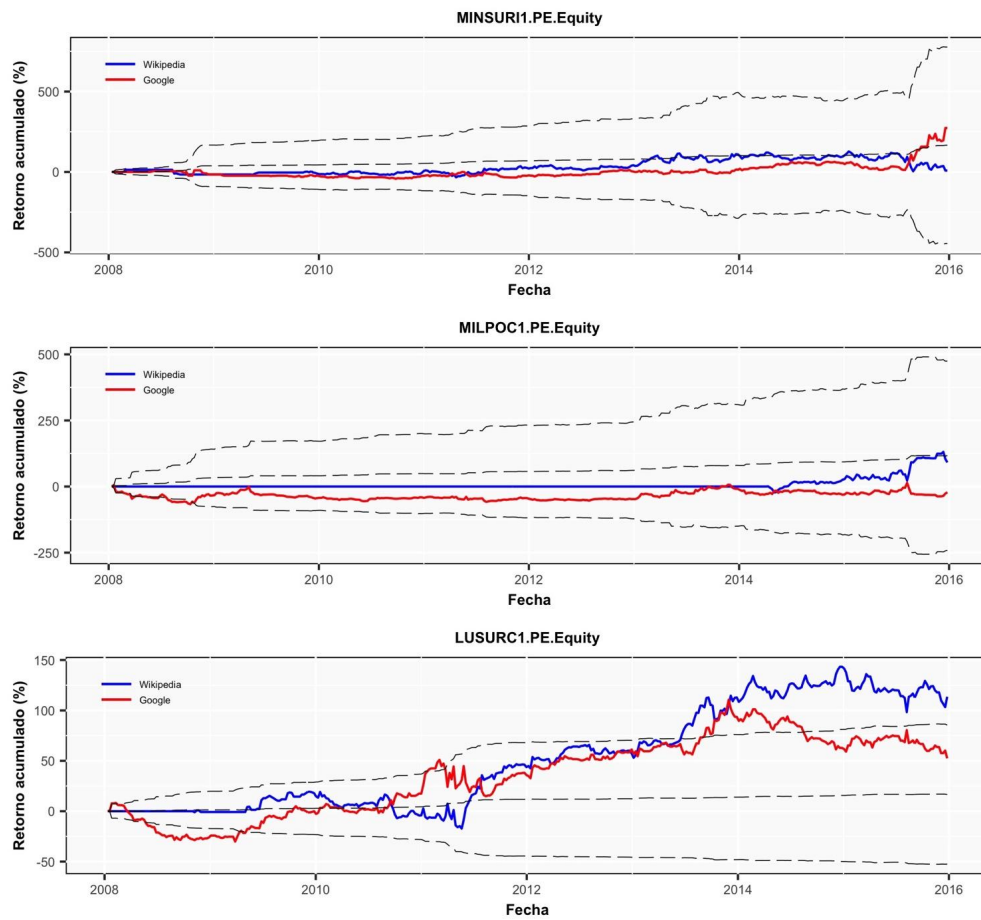
## Relación entre los patrones de búsqueda de internet y los movimientos del MILA



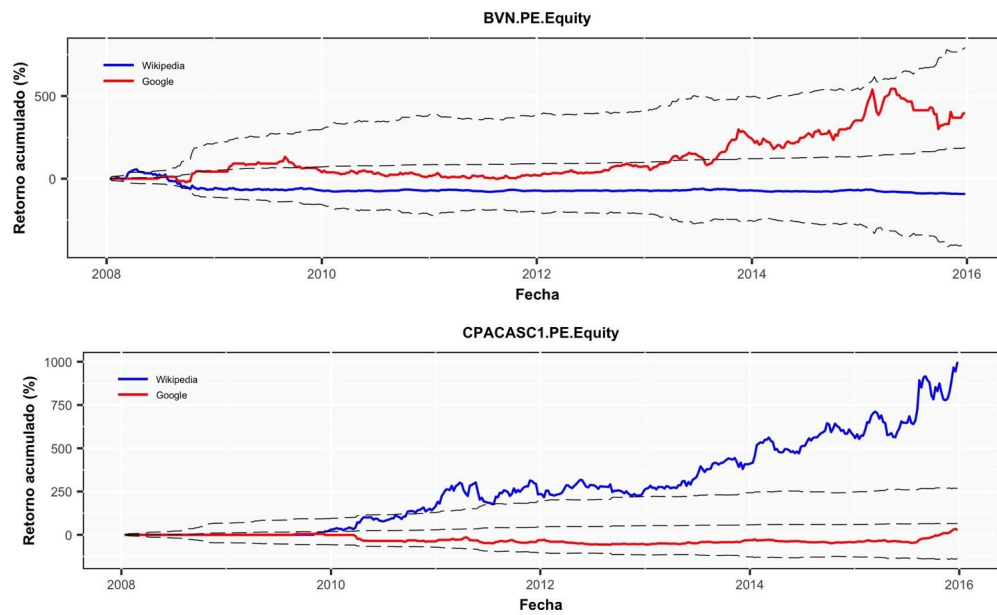
**Relación entre los patrones de búsqueda de internet y los movimientos del MILA**



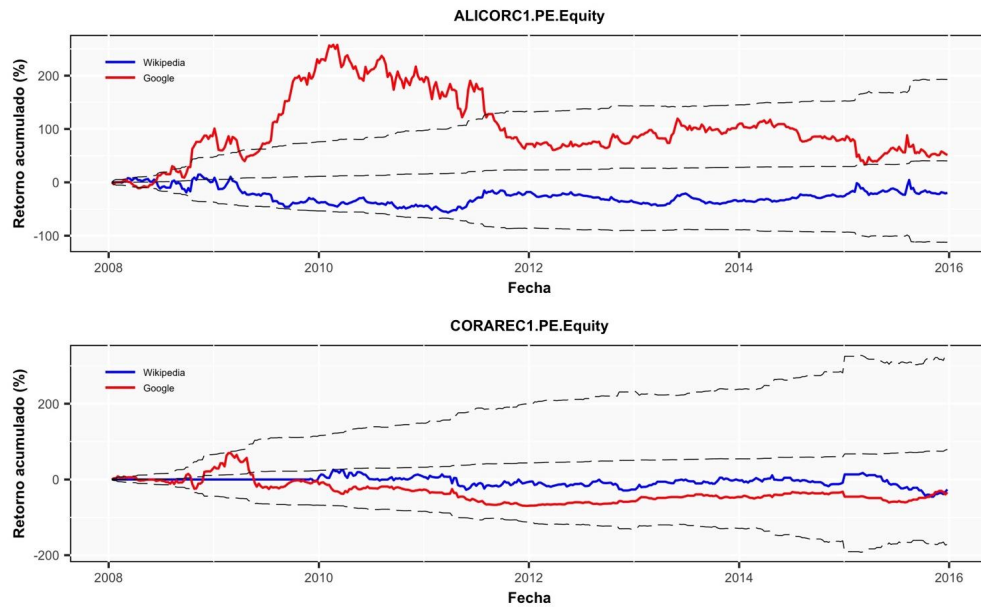
# Relación entre los patrones de búsqueda de internet y los movimientos del MILA



# Relación entre los patrones de búsqueda de internet y los movimientos del MILA



# Relación entre los patrones de búsqueda de internet y los movimientos del MILA



### ANEXO 3

*#Paquetes de datos*

**library**(reshape2)

**library**(ggplot2)

*#Precios histórico de cierre de acciones del día lunes*

precio <- **read.csv2**(file="precio.csv", head=TRUE, sep=";", dec=",")

*#Datos de Google Trends semanales*

gtrend <- **read.csv**(sep=";", file="GT.csv")

gtrend[1:8,1:4]

## semana CORFICOL.CB.Equity BOGOTA.CB.Equity GRUPOSUR.CB.Equity

## 1 12/01/08 30 24 0

## 2 19/01/08 46 34 0

## 3 26/01/08 55 29 0

## 4 2/02/08 46 41 0

## 5 9/02/08 53 36 0

## 6 16/02/08 35 39 0

## 7 23/02/08 57 34 0

## 8 1/03/08 42 43 0

### Datos de Wikipedia

leer\_wt<-function(){

*#Visitas diarias de Wikipedia en Español*

esdia<-**read.csv**(sep=";", "WT\_ES.csv")

*#Visitas diarias de Wikipedia en Inglés*

endia<-**read.csv**(sep=";", "WT\_EN.csv")

*#Unir las visitas en inglés y en español*

dia<-**merge\_Sum**(esdia, endia, "fecha", "fecha")

*#Cambio formato de fecha, diario a semanal*

semana<-**format**(**as.Date**(dia\$fecha, format="%d/%m/%y"), format="%U/%Y")

dia<-**cbind**(semana,dia)

*#Corrige la semana con el cambio de año*

for(i in 1:**nrow**(dia)){

if ((**unlist**(**strsplit**(**as.character**(dia[i,1]), split="/", fixed=TRUE)))[1]=="00")) {

dia[i,1]<-dia[i-1,1]

}

# Relación entre los patrones de búsqueda de internet y los movimientos del MILA

```

}

#Promedio semanal de visitas Wikipedia
wtrend<-aggregate(dia[,3:ncol(dia)],list(semana=dia$semana),mean)

#Ordenar cronológicamente
wtrend<-wtrend[order(as.Date(wtrend$semana, format="%U/%Y")),]

#Ajustar las fechas a las de Google y los Precios
wtrend<-wtrend[5:421,]

return(wtrend)
}

merge_Sum <- function(.df1, .df2, .id_Columns, .match_Columns){
  merged_Columns <- unique(c(names(.df1),names(.df2)))
  merged_df1 <- data.frame(matrix(nrow=nrow(.df1), ncol=length(merged_Columns)))
  names(merged_df1) <- merged_Columns
  for (column in merged_Columns){
    if(column %in% .id_Columns | !column %in% names(.df2)){
      merged_df1[, column] <- .df1[, column]
    } else if (!column %in% names(.df1)){
      merged_df1[, column] <- .df2[match(.df1[, .match_Columns],.df2[, .match_Columns]),
column]
    } else {
      df1_Values=.df1[, column]
      df2_Values=.df2[match(.df1[, .match_Columns],.df2[, .match_Columns]), column]
      df2_Values[is.na(df2_Values)] <- 0
      merged_df1[, column] <- df1_Values + df2_Values
    }
  }
  return(merged_df1)
} #importante citar

wtrend<-leer_wt()
wtrend[1:6,1:4]

## semana AESGENER.CC.Equity AGUASA.CC.Equity ANTAR.CC.Equity
## 1 01/2008 0.0000000 0 0.0000000
## 10 02/2008 0.0000000 0 0.0000000
## 19 03/2008 0.0000000 0 0.0000000
## 28 04/2008 0.3333333 0 0.0000000
## 36 05/2008 4.8571429 0 0.0000000
## 44 06/2008 3.0000000 0 0.1428571

```

## Estrategia Hipotética de Inversión

```

estrategia_hipotetica<-function(trend,ticker,deltat){

  #Lista para almacenar el retorno acumulado en cada periodo
  r<-rep(1,nrow(trend)-1)

  for(i in deltat+1:(nrow(trend)-deltat-1)){

    #Visitas en el periodo t
    now<-as.numeric(trend[[ticker]][i] )

    #Promedio de visitas periodo deltat
    previous<-0
    for(t in 1:deltat) {
      previous<-previous+trend[[ticker]][i-t]
    }
    previous<-(previous/deltat)

    #Aumento o disminución de visitas
    val<-(now-previous)

    precio_presente<-as.numeric(precio[[ticker]][i])
    precio_futuro<-as.numeric(precio[[ticker]][i+1])

    indice<-(precio_futuro/precio_presente)

    if(val>0) { #Aumentan la visitas
      # Posición corta
      r[i]<-r[i-1]/ indice}

    if(val<0) {#Disminuyen las visitas
      # Posición Larga
      r[i]<-r[i-1]* indice}

    if(val==0) { # No toma posición
      r[i]<-r[i-1]}
  }
  r <- (r-1)*100 #Retorno en términos porcentuales
  return(r)
}

ticker<-"EXITO.CB.Equity"
R_Wikipedia<- estrategia_hipotetica(wtrend,ticker,3)
R_Google<- estrategia_hipotetica(gtrend,ticker,1)

```



## Estrategia Aleatoria de Inversión

```

estrategia_aleatoria<-function(trend,ticker,deltat,iteraciones){

  # Matriz para almacenar el retorno acumulado
  r<-matrix(data=1,nrow=nrow(trend)-1,ncol=iteraciones)

  for(j in 1:iteraciones){

    for(i in deltat+1:(nrow(trend)-deltat-1)) {

      precio_presente <- as.numeric(precio[[ticker]][i])

      precio_futuro <- as.numeric(precio[[ticker]][i+1])

      indice_precio <- precio_futuro/precio_presente

      val<-sample(0:1,1) #Generación número aleatorio (0 o 1)

      if(val==0) { # Posición corta
        r[i,j] <- r[i-1,j] / indice_precio }

      if(val==1) { # Posición larga
        r[i,j] <- r[i-1,j] * indice_precio }
      }
    }
  }

  r <- (r-1)*100

  #Promedio retorno acumulado de cada periodo (fila)
  media<-apply(r,1,mean)
  #Desviación estandar del retorno acumulado en cada periodo (fila)
  desv<-apply(r,1,sd)

  #Resultado:fecha, media, desviación estandar
  fecha<-as.Date(trend[1:nrow(trend)-1,1],"%d/%m/%y")
  resultado<-data.frame(cbind(media,desv))
  resultado<-data.frame(cbind(fecha,resultado))

  return(resultado) }

aleatoria <- estrategia_aleatoria(gtrend,ticker,1,10000)
datos<-cbind(aleatoria,R_Wikipedia,R_Google)
head(datos)

```

```
##      fecha      media      desv R_Wikipedia R_Google
## 1 2008-01-12 0.00000000 0.000000 0.0000000 0.000000
## 2 2008-01-19 0.07767448 5.348274 0.0000000 5.491329
## 3 2008-01-26 0.07767448 5.348274 0.0000000 5.491329
## 4 2008-02-02 0.07266029 5.398216 0.6896552 6.218856
## 5 2008-02-09 0.05208495 6.114944 -2.1447721 9.295540
## 6 2008-02-16 0.05544892 6.116975 -2.0134228 9.442245
```

## Gráfico Estrategia Hipotética y Aleatoria

```
grafico<-function(datos){
  df<-data.frame(datos)

  ggplot(df, aes(x=fecha)) +
    geom_line(aes(y=R_Wikipedia,colour="Wikipedia"))+
    geom_line(aes(y=R_Google,colour="Google"))+
    scale_colour_manual("",
      breaks = c("Wikipedia", "Google"),
      values = c("Wikipedia"="blue", "Google"="red"))+
    geom_line(aes(y=media-desv),linetype=5,size=0.2)+
    geom_line(aes(y=media+desv),linetype=5,size=0.2)+
    geom_line(aes(y=media),linetype=5, size=0.2)+
    labs(title =ticker, x = "Fecha", y = "Retorno acumulado (%)")+
    theme(plot.title = element_text(size=7,face="bold"),
      axis.title = element_text(size=7,face="bold"),
      axis.text=element_text(size=7),
      panel.background = element_rect(fill = 'gray98',colour = "black"),
      legend.text=element_text(size=7),
      legend.position=c(0.08,0.9),
      legend.background = element_rect(fill = NA),
      legend.key = element_rect(fill = NA, colour=NA),
      legend.key.height=unit(0.5,"line")
    )
}

library(ggplot2)
grafico(datos)
```

## Tabla de resultados

```
df<-data.frame()
for(n in 2:ncol(gtrend)){
```

## Relación entre los patrones de búsqueda de internet y los movimientos del MILA

```
ticker<-colnames(gtrend)[n]
for(deltat in 1:6){
  R.Wikipedia<-estrategia_hipotetica(wtrend,ticker,deltat)[416]
  R.Google<-estrategia_hipotetica(gtrend,ticker,deltat)[416]
  aleatoria<-estrategia_aleatoria(gtrend,ticker,deltat,500)[416,2:3]
  row<-cbind(ticker,deltat,R.Google,R.Wikipedia,aleatoria)
  df<-rbind(df,row) } }

#Dividir el ticker para obtener el país
df<-cbind(colsplit(df$ticker,pattern="[.]",names=c("a","pais","b")),df)
df[,1]<-NULL
df[,2]<-NULL

df$deltat<- as.character(df$deltat)

#Retorno aritméticos a logarítmicos
df$R.Wikipedia<-log(1+df$R.Wikipedia/100)*100
df$R.Google<-log(1+df$R.Google/100)*100
```

## Gráfico de Cajas

```
dat <- aggregate(cbind(R.Wikipedia,R.Google) ~ deltat, df, mean)
dat$R.Wikipedia<-round(dat$R.Wikipedia,2)
dat$R.Google<-round(dat$R.Google,2)
nom<-c(CC="Chile",MM="México",CB="Colombia",PE="Perú")

ggplot(df, aes(x=deltat, y=R.Google) )+
  geom_boxplot()+
  stat_summary(fun.y=mean, colour="darkred", geom="point", shape=18, size=3) +
  # facet_wrap(~pais,scales="free_y",labeller=labeler(pais=nom )) +
  geom_text(data = dat, aes(label = R.Google, y =R.Google - 20))+
  labs( x = "Delta t", y = "Retorno Acumulado Google(%)" )+
  theme(plot.title = element_text(size=15,face="bold"),
        axis.title = element_text(size=15,face="bold"),
        axis.text=element_text(size=15))+
  coord_flip()
```

## Gráfico de Barras

```
df$Goog_desv<-df$R.Google/df$desv
df$Wiki_desv<-df$R.Wikipedia/df$desv
dat<- aggregate(cbind(Wiki_desv,Goog_desv)~ticker+pais, df, mean)

dat<-melt(dat,id=c("ticker","pais"))
dat.co<-subset(dat,pais=="MM")

ggplot(dat.co, aes(x = reorder(ticker, value),y=value, fill=variable))+
  #facet_grid(. ~ pais,scales="free",labeller=labeler(pais=nom))+
  geom_bar(width=0.7,stat="identity",position=position_dodge(width=0.7))+
  coord_flip()+
  labs(x="", y = "Retorno acumulado (%)")+
  scale_fill_manual(values=c("red", "blue"),
                    name="",
                    breaks=c("R.Google", "R.Wikipedia"),
                    labels=c("Google", "Wikipedia"))
```











